

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 6, 2004, 04:55:28 ; Search time 1265 Seconds
(without alignments)
4458.371 Million cell updates/sec

Title: US-09-882-509-3

Perfect score: 1245

Sequence: 1 attgttgacgtgagtgcgt.....ataaccctaagacaataa 1245

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 2941586 seqs, 2264995651 residues

Total number of hits satisfying chosen parameters: 5883172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/FCI_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1224.2	98.3	1245	10	US-09-940-235-1
2	1222.6	98.2	1377	10	US-09-940-235-5
3	1222.6	98.2	1782	10	US-09-940-235-11
4	1209.8	97.2	1327	10	US-09-940-235-6
5	1127.8	90.6	1661	10	US-09-940-235-10
6	1124.6	90.3	2096	10	US-09-940-235-12
7	1111.8	89.3	1541	10	US-09-940-235-9
8	43.4	3.5	556	16	US-10-369-493-29120
9	42.4	3.4	5504	13	US-10-221-613-270
10	42.4	3.4	6071	15	US-10-311-455-340
11	42.4	3.4	6071	15	US-10-240-485-36
12	42.4	3.4	3673778	15	US-10-312-841-1
13	41.8	3.4	6392	15	US-10-311-455-658
14	41.8	3.4	6392	15	US-10-240-485-60

C 15	41.6	3.3	6283	15	US-10-311-455-61	Sequence 61, Appl
C 16	40.8	3.3	1794	13	US-10-282-122A-40523	Sequence 40523, A
C 17	40.8	3.3	9504	15	US-10-240-453-281	Sequence 281, App
C 18	40.4	3.2	45	10	US-09-940-235-27	Sequence 27, Appl
C 19	40.4	3.2	9084	15	US-10-311-455-1578	Sequence 1578, Ap
C 20	40.2	3.2	53	10	US-09-940-235-14	Sequence 14, Appl
C 21	40.2	3.2	2000	9	US-09-938-842A-3923	Sequence 3923, Ap
C 22	40.2	3.2	2000	11	US-09-938-842A-3923	Sequence 3923, Ap
C 23	40.2	3.2	6407	15	US-10-311-455-916	Sequence 916, App
C 24	40.2	3.2	8234	15	US-10-311-455-1622	Sequence 1622, Ap
C 25	40.2	3.2	3673778	15	US-10-312-841-2	Sequence 2, Appli
C 26	39.6	3.2	13326	15	US-10-311-455-1686	Sequence 1686, Ap
C 27	39.4	3.2	19734	15	US-10-311-455-1906	Sequence 1906, Ap
C 28	39.4	3.2	580073	15	US-10-205-220-1	Sequence 1, Appli
C 29	38.8	3.1	5690	13	US-10-240-454-26	Sequence 26, Appl
C 30	38.8	3.1	5690	15	US-10-311-455-1298	Sequence 1298, Ap
C 31	38.8	3.1	5690	17	US-10-257-166-110	Sequence 110, App
C 32	38.8	3.1	16228	13	US-10-221-613-387	Sequence 387, App
C 33	38.6	3.1	640681	9	US-09-790-988-1	Sequence 1, Appli
C 34	38.4	3.1	1079	9	US-09-070-927A-674	Sequence 674, App
C 35	38.4	3.1	1890	13	US-10-282-122A-21111	Sequence 2111, A
C 36	38.2	3.1	609	10	US-09-769-736-49	Sequence 49, Appl
C 37	38.2	3.1	2232	15	US-10-087-464-45	Sequence 45, Appl
C 38	38	3.1	5514	15	US-10-311-455-103	Sequence 103, App
C 39	38	3.1	14095	15	US-10-311-455-450	Sequence 450, App
C 40	37.8	3.0	6104	13	US-10-221-714A-17	Sequence 17, Appl
C 41	37.8	3.0	6104	15	US-10-311-455-269	Sequence 269, App
C 42	37.8	3.0	8244	13	US-10-221-714A-117	Sequence 117, App
C 43	37.8	3.0	8244	15	US-10-311-455-365	Sequence 365, App
C 44	37.4	3.0	588	13	US-10-027-632-265227	Sequence 265227,
C 45	37.4	3.0	588	16	US-10-027-632-265227	Sequence 265227,

ALIGNMENTS

RESULT 1

US-09-940-235-1
Sequence 1, Application US/09940235
Publication No. US2003005921A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammar
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/09/940,235
CURRENT FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1245
TYPE: DNA
ORGANISM: Streptococcus equisimilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)....(1242)
US-09-940-235-1

Query Match 98.3%; Score 1224.2; DB 10; Length 1245;
Best Local Similarity 99.0%; Pred No. 0;
Matches 1232; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```
Qy 1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATTAGTTGTT 60
Db 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATTAGTTGTT 60
Qy 61 AGCGTTGCTGTAATGTTGAGGGAACGAATCAAGACATTAAGTCTTAAATTTTGAATTT 120
Db 61 AGCGTTGCTGTAATGTTGAGGGAACGAATCAAGACATTAAGTCTTAAATTTTGAATTT 120
Qy 121 GACCTAACATCACCACCTGCTCATGAGGGAAGACAGACGAGGCTTAAGTCCAAATCA 180
Db 121 GATCTAACATCACCACCTGCTCATGAGGGAAGACAGACGAGGCTTAAGTCCAAATCA 180
Qy 181 AAACCAATTTGCTACTGATGAGTGGCGGATGCGCATAAATCTGAAAAGCTGACTTACTA 240
Db 181 AAACCAATTTGCTACTGATGAGTGGCGGATGCGCATAAATCTGAAAAGCTGACTTACTA 240
Qy 241 AAGCTATTCAAGAACCAATTGATCGCTAAACGTCACAGTAAACGACGACTACTTTGAGGTC 300
Db 241 AAGCTATTCAAGAACCAATTGATCGCTAAACGTCACAGTAAACGACGACTACTTTGAGGTC 300
Qy 301 ATTGATTTTGAAGCGATGCAACATTAAGTCTGATGAAACGCGAAGTCTACTTTGCTGAC 360
Db 301 ATTGATTTTGAAGCGATGCAACATTAAGTCTGATGAAACGCGAAGTCTACTTTGCTGAC 360
Qy 361 AAAGATGGTTGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 420
Db 361 AAAGATGGTTGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 420
Qy 421 GTGCGGCTTAGACCATATAAGAAAACCAATACAAAATCAAGGAAATCTGTGATGTG 480
Db 421 GTGCGGCTTAGACCATATAAGAAAACCAATACAAAATCAAGGAAATCTGTGATGTG 480
Qy 481 GAATATCTCTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGCAGGCTCAAA 540
Db 481 GAATATCTCTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGCAGGCTCAAA 540
Qy 541 GATACTAAGCTATTGAAAACTAGTCTATCGGTGACCCATCATCTCAAGAAATTAATA 600
Db 541 GATACTAAGCTATTGAAAACTAGTCTATCGGTGACCCATCATCTCAAGAAATTAATA 600
Qy 601 GCTCAAGCACAAGCATTTTAAACAAACCCAGCTTATAGGATCAATATAAAAAATCTGGTCTG 660
Db 601 GCTCAAGCACAAGCATTTTAAACAAACCCAGCTTATAGGATCAATATAAAAAATCTGGTCTG 660
Qy 661 TCCTCAATCTGCTCAATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
Db 661 TCCTCAATCTGCTCAATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
Qy 721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATATAAAAAATCTGGTCTG 780
Db 721 TTTACTTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAATATAAAAAATCTGGTCTG 780
Qy 781 AATGAAGAAATAACACATGACCTCTCTGAGAAATATTACGTCCTTAAARAGGG 840
Db 781 AATGAAGAAATAACACATGACCTCTCTGAGAAATATTACGTCCTTAAARAGGG 840
Qy 841 GAAAAGCCGTATGATCCCTTTGATGCGAGTCACTTTGAAAATCTGTTCAACATCAATACGTT 900
Db 841 GAAAAGCCGTATGATCCCTTTGATGCGAGTCACTTTGAAAATCTGTTCAACATCAATACGTT 900
Qy 901 GATGTCACACCAAGCAATTTGCTTAAAGGAGAGGCTCTTAAACAGTACGGAACGTAAC 960
Db 901 GATGTCACACCAAGCAATTTGCTTAAAGGAGAGGCTCTTAAACAGTACGGAACGTAAC 960
Qy 961 TTAGCTTTACAGATTTTATACGATCCCTGATAGGCTAAATCTCTCAACATCTC 1020
Db 961 TTAGCTTTACAGATTTTATACGATCCCTGATAGGCTAAATCTCTCAACATCTC 1020
Qy 1021 GATGCTTTTGGTATTATGGAATACTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1080
Db 1021 GATGCTTTTGGTATTATGGAATACTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1080
```

```
Qy 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
Db 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
Qy 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAAAGACGAGAAAGTTTACAGTACTG 1200
Db 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAAAGACGAGAAAGTTTACAGTACTG 1200
Qy 1201 CGTTATACAGGACACTATATACCTGATTAACCTTAACGACAAATAA 1245
Db 1201 CGTTATACAGGACACTATATACCTGATTAACCTTAACGACAAATAA 1245

RESULT 2
US-09-940-235-5
; Sequence 5, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-5

Query Match 98.2%; Score 1222.6; DB 10; Length 1377;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATTAGTTGTT 60
Db 133 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATTGGTTGTT 192
Qy 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATTT 120
Db 193 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATTT 252
Qy 121 GACCTAACATCAGACCTGCTCATGAGGAAAGACAGACGAGGCTTAAAGTCCAAATCA 180
Db 253 GATCTAACATCAGACCTGCTCATGAGGAAAGACAGACGAGGCTTAAAGTCCAAATCA 312
Qy 181 AAACCAATTTGCTACTGATGAGTGGCGGATGCGCATAAATCTGAAAAGCTGACTTACTA 240
Db 313 AAACCAATTTGCTACTGATGAGTGGCGGATGCGCATAAATCTGAAAAGCTGACTTACTA 372
Qy 241 AAGGCTATTCAAGAACCAATTGATCGTAAAGTCCACAGTAAACGACGACTACTTTGAGGTC 300
Db 373 AAGGCTATTCAAGAACCAATTGATCGTAAAGTCCACAGTAAACGACGACTACTTTGAGGTC 432
Qy 301 ATTGATTTTGAAGCGATGCAACATTAAGTCTGATGAAACGCGAAGGCTTACTTTGCTGAC 360
Db 433 ATTGATTTTGAAGCGATGCAACATTAAGTCTGATGAAACGCGAAGGCTTACTTTGCTGAC 492
Qy 361 AAAGATGGTTGGTAACTTTGCCGACCCCAACCTGTCCAAGAAATTTTGTAAAGCGGACAT 420
```

```

Db 493 AAAGATGGTTCGGTAACCTTGGCGACCAACCTGTCCAGAAATTTTGTCTAAGCGACAT 552
QY 421 GTGCGGTTAGACATATAAGAAACCAACCAATACAAATCAAGCGAAATCTTTGATGTG 480
Db 553 GTGCGGTTAGACATATAAGAAACCAACCAATACAAATCAAGCGAAATCTTTGATGTG 612
QY 481 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 540
Db 613 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 672
QY 541 GATATAGTGTATGAAACCACTAGTATCGGTGACCAACATCACTCTCAAGAAATTA 600
Db 673 GATATAGTGTATGAAACCACTAGTATCGGTGACCAACATCACTCTCAAGAAATTA 732
QY 601 GCTCAGACCAAGCAATTTTAAACCAACCCACCCAGGCTATAGATTTTATGACGAGT 660
Db 733 GCTCAGACCAAGCAATTTTAAACCAACCCACCCAGGCTATAGATTTTATGACGAGT 792
QY 661 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
Db 793 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 852
QY 721 TTTACTTACATGTCATGCAAAATCGGGAACAGCTTTATGAGATCAATAAAAATCTGCTG 780
Db 853 TTTACTTACATGTCATGCAAAATCGGGAACAGCTTTATGAGATCAATAAAAATCTGCTG 912
QY 781 AATGAGAAATTAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
Db 913 AATGAGAAATTAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 972
QY 841 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCAATCAATACGTT 900
Db 973 GAAAGCCGATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCAATCAATACGTT 1032
QY 901 GATGTCACACCAAGAAATGCTTAAAGAGGAGAGCTCTTAAAGCTAGCGAAACGTAAC 960
Db 1033 GATGTCACACCAAGAAATGCTTAAAGAGGAGAGCTCTTAAAGCTAGCGAAACGTAAC 1092
QY 961 TTAGACTTCAGAGATTTATACGATCCTCGTGAAGGCTAAAGCTCTCTACACAAATCTC 1020
Db 1093 TTAGACTTCAGAGATTTATACGATCCTCGTGAAGGCTAAAGCTCTCTACACAAATCTC 1152
QY 1021 GATGCTTTTGGTATTATGCACTATACCTTAACTGGAAAGTAGAGGATATACGATGAC 1080
Db 1153 GATGCTTTTGGTATTATGCACTATACCTTAACTGGAAAGTAGAGGATATACGATGAC 1212
QY 1081 ACCAACCGTATCATACCGTTTATGCGCAAGGACCCGAGGAGAGATGCTAGCTAT 1140
Db 1213 ACCAACCGTATCATACCGTTTATGCGCAAGGACCCGAGGAGAGATGCTAGCTAT 1272
QY 1141 CATTTAGCCTTATGATAAAGATCGTTATACCGAAGAGAGACGAGAGTTTACAGTACCTG 1200
Db 1273 CATTTAGCCTTATGATAAAGATCGTTATACCGAAGAGAGACGAGAGTTTACAGTACCTG 1332
QY 1201 CGTTATACAGGACCTTATACCTGATAACCCCTTAACGACAAATPAA 1245
Db 1333 CGTTATACAGGACCTTATACCTGATAACCCCTTAACGACAAATPAA 1377

```

RESULT 3
 US-09-940-235-11
 ; Sequence 11, Application US/09940235
 ; Publication No. US20030059921A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sanni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammarra
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE

; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; CURRENT FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11
 ; LENGTH: 1782
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Hybrid cassette
 ; US-09-940-235-11

Query Match 98.2%; Score 1222.6; DB 10; Length 1782;
 Best Local Similarity 98.9%; Pred. No. 0; Mismatches 14; Indels 0; Gaps 0;
 Matches 1231; Conservative 0;

QY	1	ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATAGTTGTT	60
Db	538	ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATAGTTGTT	597
QY	61	AGCGTTGCTGGTACTTGTGAGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATTT	120
Db	598	AGCGTTGCTGGTACTTGTGAGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAATTT	657
QY	121	GACCTAAATCAACGACCTGCTCATGGAGGAAGACAGACAGCGCTTAAGTCCAAATCA	180
Db	658	GATCTAAATCAACGACCTGCTCATGGAGGAAGACAGACAGCGCTTAAGTCCAAATCA	717
QY	181	AAACCAATTTGCTACTGATAGTGGCGGATGCCACATATAAATCTGAAAAGCTGCTTACTA	240
Db	718	AAACCAATTTGCTACTGATAGTGGCGGATGCCACATATAAATCTGAAAAGCTGCTTACTA	777
QY	241	AGGCTATTCAAGAACAAATGATCGCTTAACGCTCCACAGTAAACGACACTACTTTGAGTTC	300
Db	778	AGGCTATTCAAGAACAAATGATCGCTTAACGCTCCACAGTAAACGACACTACTTTGAGTTC	837
QY	301	ATTGATTTTGCAGCGCATCAACCAATTAATCGATCGAAACGCGAAGTCTTACTTTGCTGAC	360
Db	838	ATTGATTTTGCAGCGCATCAACCAATTAATCGATCGAAACGCGAAGTCTTACTTTGCTGAC	897
QY	361	AAAGATGGTTCGGTAACTTTGCGGACCCCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT	420
Db	898	AAAGATGGTTCGGTAACTTTGCGGACCCCAACCTGTCCAAAGAAATTTTGTCTAAGCGGACAT	957
QY	421	GTGCGGTTAGACCATATAAGAAACCAATACAAATCAAGAAATCTGTTGATGTG	480
Db	958	GTGCGGTTAGACCATATAAGAAACCAATACAAATCAAGAAATCTGTTGATGTG	1017
QY	481	GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA	540
Db	1018	GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA	1077
QY	541	GATATAGCTATTGAAACCACTAGCTATCGGTGACCAACATCAATCTCAAGAAATTA 600	
Db	1078	GATATAGCTATTGAAACCACTAGCTATCGGTGACCAACATCAATCTCAAGAAATTA 1137	
QY	601	GCTCAGACCAAGCAATTTTAAACCAACCCACCCAGGCTATACGATTTTATGACGAGT 660	
Db	1138	GCTCAGACCAAGCAATTTTAAACCAACCCACCCAGGCTATACGATTTTATGACGAGT 1197	
QY	661	TCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720	
Db	1198	TCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 1257	
QY	721	TTTACTTACCATGTCAAAATTCGGGAACAGCTTTATGAGATCAATAAAAATCTGCTG 780	

Db 1258 TTTACTTACCGTGTAAATAATCGGAAACAAGCTTATAGGATCAATAAAAAATCTGGTCTG 1317
QY 781 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG 840
Db 1318 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGG 1377
QY 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 900
Db 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 1437
QY 901 GATGTCAACACCAAGGAATGCTTAAAGCGAGCGAGCTCTTAAACAGCTAGCGAAGTAA 960
Db 1438 GATGTCAACACCAAGGAATGCTTAAAGCGAGCGAGCTCTTAAACAGCTAGCGAAGTAA 1497
QY 961 TTAGACTTTCAGAGATTTATACGATCCTCGTATAGGCTTAAACTCTCTACAACTCTC 1020
Db 1498 TTAGACTTTCAGAGATTTATACGATCCTCGTATAGGCTTAAACTCTCTACAACTCTC 1557
QY 1021 GATGCTTTGGTATTTATGAGTATACCTTAACTGAAAGTAGAGGATTAATCAGATGAC 1080
Db 1558 GATGCTTTGGTATTTATGAGTATACCTTAACTGAAAGTAGAGGATTAATCAGATGAC 1617
QY 1081 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
Db 1618 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1677
QY 1141 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAAACGAGAGTTTACAGCTACCTG 1200
Db 1678 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAAACGAGAGTTTACAGCTACCTG 1737
QY 1201 CGTTATACAGGACCACTTATACCTGATTAACCTTAACGCAAAATAA 1245
Db 1738 CGTTATACAGGACCACTTATACCTGATTAACCTTAACGCAAAATAA 1782

RESULT 4

US-09-940-235-6
; Sequence 6, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR FILING DATE: 1999-12-23
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-6

Query Match 97.2%; Score 1209.8; DB 10; Length 1327;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1223; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACGCAATAGTTGTT 60

Db 83 ATAGCTGGTCTCGTAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATGCTTGT 142
QY 61 AGCGTCTGCTGCTACTGTTGAGGGACGAATCAAGCATTAGTCTTAAATTTTGAATTT 120
Db 143 AGCGTCTGCTGCTACTGTTGAGGGACGAATCAAGCATTAGTCTTAAATTTTGAATTT 202
QY 121 GACCTAAACATCACCACCTGCTCATGAGGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 180
Db 203 GATCTAACATCAGCACTGCTCATGAGGAGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 262
QY 181 AAACCATTTGCTACTGATAGTGGCGGATGCAATAACTTGAAGAAAGCTGACTTACTA 240
Db 263 AAACCATTTGCTACTGATAGTGGCGGATGCAATAACTTGAAGAAAGCTGACTTACTA 322
QY 241 AAGGCTATTCAAGAAACAATTGATCGCTAAAGTCCACAGTAAAGCAAGCTACTTTGAGT 300
Db 323 AAGGCTATTCAAGAAACAATTGATCGCTAAAGTCCACAGTAAAGCAAGCTACTTTGAGT 382
QY 301 ATTGATTTTGAAGCGATGCAACCAATTTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 360
Db 383 ATTGATTTTGAAGCGATGCAACCAATTTACTGATCGAAACCGCAAGGCTTACTTTGCTGAC 442
QY 361 AAGATGCTTGGTTCGTTAACTTGGCGACCACTGTCAGAAATTTTGTGAAGCGGAT 420
Db 443 AAGATGCTTGGTTCGTTAACTTGGCGACCACTGTCAGAAATTTTGTGAAGCGGAT 502
QY 421 GTGCGCTTAGACCATATAAGAAACCAATTAATAAATCAAGGAAATCTGTTGATGTG 480
Db 503 GTGCGCTTAGACCATATAAGAAACCAATTAATAAATCAAGGAAATCTGTTGATGTG 562
QY 481 GAATATCTGTACAGTTTACTTCCCTTAAACCTGTATGATGACGATTTTCAAGCAAGTCTCAAA 540
Db 563 GAATATCTGTACAGTTTACTTCCCTTAAACCTGTATGATGACGATTTTCAAGCAAGTCTCAAA 622
QY 541 GATATTAAGCTTATGAAACAATGATGATCGTGGTACACCATCATCTCAAGAAATTA 600
Db 623 GATATTAAGCTTATGAAACAATGATGATCGTGGTACACCATCATCTCAAGAAATTA 682
QY 601 GCTCAAGCAACAAAGCATTTTAAACAAACCCAGGCTATAGATTTTATGAACGTTGAC 660
Db 683 GCTCAAGCAACAAAGCATTTTAAACAAACCCAGGCTATAGATTTTATGAACGTTGAC 742
QY 661 TCCTCAATGCTGCTCATGATCAATGATGATTTCCGTACGATTTTACCAATGGATCAAGAG 720
Db 743 TCCTCAATGCTGCTCATGATCAATGATGATTTCCGTACGATTTTACCAATGGATCAAGAG 802
QY 721 TTTACTTACCATGTCRAAATCGGACACAGCTTATGAGATCAATAAATCTGCTG 780
Db 803 TTTACTTACCATGTCRAAATCGGACACAGCTTATGAGATCAATAAATCTGCTG 862
QY 781 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
Db 863 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 922
QY 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 900
Db 923 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 982
QY 901 GATGTCAACCAACCAAGGAATGCTTAAAGCGAGCGAGCTCTTAAACAGCTAGCGAAGTAA 960
Db 983 GATGTCAACCAACCAAGGAATGCTTAAAGCGAGCGAGCTCTTAAACAGCTAGCGAAGTAA 1042
QY 961 TTAGACTTTCAGAGATTTATACGATCCTCGTATAGGCTTAAACTCTCTACAACTCTC 1020
Db 1043 TTAGACTTTCAGAGATTTATACGATCCTCGTATAGGCTTAAACTCTCTACAACTCTC 1102
QY 1021 GATGCTTTGGTATTTATGAGTATACCTTAACTGGAAGTAGAGGATTAATCAGATGAC 1080
Db 1103 GATGCTTTGGTATTTATGAGTATACCTTAACTGGAAGTAGAGGATTAATCAGATGAC 1162
QY 1081 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
Db 1163 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1222

Qy	1141	CATTAGCCTTATGATAAAGATCGTTATACCGAAGAAACGAGAAGTTTACAGCTACCTG	1200
Db	1223	CATTAGCCTTATGATAAAGATCGTTATACCGAAGAAACGAGAAGTTTACAGCTACCTG	1282
Qy	1201	CGTTATACGGGACCTTATACCTGATAACCTTAACGACAAATAA	1245
Db	1283	CGTTATACGGGACCTTATACCTGATAACCTTAACGACAAATAA	1327

RESULT 5

```

US-09-940-235-10
; Sequence 10, Application US/09940235
; Publication NO. US2003059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1661
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-10

```

543	AAAGATGGTTTCGGTAACCTTGGCGACCCAACTGTGTCGAAGAAATTTTTTGTCTAAGCGGACAT	602
421	GTGCGGTTTAGACCATATAAAGAAAAACCAATCAAAAAACAAGCGAAATCTGTGTATGTG	480
603	GTGCGGTTTAGACCATATAAAGAAAAACCAATACAAAAACAAGCGAAATCTGTGTATGTG	662
481	GAATATACCTGTACAGTTTACTCCCTTAAACCCTGTATGACGATTTAGACAGGTCCTCAAA	540
663	GAATATACCTGTACAGTTTACTCCCTTAAACCCTGTATGACGATTTAGACAGGTCCTCAAA	722
541	GATACCTAAGCTATTGAAAAACCTAGCTATCTCGGTGACACCATCACATCTCAAGAAATTTACTA	600
723	GATACCTAAGCTATTGAAAAACCTAGCTATCTCGGTGACACCATCACATCTCAAGAAATTTACTA	782
601	GCTCAAGCACAAGCAATTTTAAACAAAACCCACCCAGGCTATACGATTTTATGAACGTGAC	660
783	GCTCAAGCACAAGCAATTTTAAACAAAACCCACCCAGGCTATACGATTTTATGAACGTGAC	842
661	TCCCTCAATCGTCACTCATGACATGACATTTTCCGTCACGATTTTACCATGGATCAAGAG	720
843	TCCCTCAATCGTCACTCATGACATGACATTTTCCGTCACGATTTTACCATGGATCAAGAG	902
721	TTTTACTTACCATGTCAAAAATCGGGAAACAAGCTTTATGAGTCAATATAAAAAATCTGGTCTG	780
903	TTTTACTTACCGTGTTAATAATCGGGAAACAAGCTTTATGAGTCAATATAAAAAATCTGGTCTG	962
781	AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAAGGG	840
963	AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAAGGG	1022
841	GAAGAAGCGTATGATCCCTTTGATCGAGTCACTTTGAAACTGTTTACCACTCAAAATACGTT	900
1023	GAAGAAGCGTATGATCCCTTTGATCGAGTCACTTTGAAACTGTTTACCACTCAAAATACGTT	1082
901	GATGTCACACCAAGAAATGGCTTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAAACGTAAC	960
1083	GATGTCGATACCAACGAATGGCTTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAACGTAAC	1142
961	TTAGACTTCAGAGATTTTATACGATCTCTCGTATAGAGGCTTAACTCTTCAACAAATCTC	1020
1143	TTAGACTTCAGAGATTTTATACGATCTCTCGTATAGAGGCTTAACTCTTCAACAAATCTC	1202
1021	GATGCTTTTGGTATTATGACTATACCTTAACTCGGAAGATGAGGATTAATCAGCATGAC	1080
1203	GATGCTTTTGGTATTATGACTATACCTTAACTCGGAAGATGAGGATTAATCAGCATGAC	1262
1081	ACCAACCGGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT	1140
1263	ACCAACCGGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAAATGCTAGCTAT	1322
1141	CATTTAGCCTATGAT	1155
1323	CATTTAGCCGGTGGT	1337

RESULT 6

US-09-940-235-12
 / Sequence 12, Application US/09940235
 / Publication No. US2003005992A1
 / GENERAL INFORMATION:
 / APPLICANT: Kumar, Rajesh
 / APPLICANT: Sahni, Girish
 / APPLICANT: Roy, Chait
 / APPLICANT: Rajagopal, Kammarra
 / APPLICANT: Nihalani, Deepak
 / APPLICANT: Sundaram, Vasudha
 / APPLICANT: Yadav, Mahavir
 / TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 / TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 / TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 / TITLE OF INVENTION: PROTEIN
 / FILE REFERENCE: 07064-009002
 / CURRENT APPLICATION NUMBER: US/09/940-235

```

; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-12

Query Match      30.3%; Score 1124.6; DB 10; Length 2096;
Best Local Similarity 98.4%; Pred. No. 1.3e-302;
Matches 1136; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATAGTTGTT 60
Db 588 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATAGTTGTT 647
QY 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 120
Db 648 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 707
QY 121 GACCTAACATCAGACCTGCTCATGGAGGAAAGACAGACAGCAAGCTTAAGTCCAAATCA 180
Db 708 GATCTAACATCAGACCTGCTCATGGAGGAAAGACAGACAGCAAGCTTAAGTCCAAATCA 767
QY 181 AAACCAATTTGCTACTGATAGTGGCGGATGCGACATAACTTCAAAAGCTGACTACTA 240
Db 768 AAACCAATTTGCTACTGATAGTGGCGGATGCTCATAACTTCAAAAGCTGACTACTA 827
QY 241 AAGCTATTCAAGAAACAATTGATCGTTAAGTCCACAGTAAAGACGACTACTTTAGGTC 300
Db 828 AAGCTATTCAAGAAACAATTGATCGTTAAGTCCACAGTAAAGACGACTACTTTAGGTC 887
QY 301 ATTGATTTTCCAGCGATGCAACCATCTACTGATCGAAACGCAAGGCTACTTTGCTGAC 360
Db 888 ATTGATTTTCCAGCGATGCAACCATCTACTGATCGAAACGCAAGGCTACTTTGCTGAC 947
QY 361 AAAGATGTTTCGGTAACTCTTCCGACCCCAACCTGTCACAGAAATTTTTCGTAAGCGGACAT 420
Db 948 AAAGATGTTTCGGTAACTCTTCCGACCCCAACCTGTCACAGAAATTTTTCGTAAGCGGACAT 1007
QY 421 GTGGCGTTTACACATATAAGAAAAAACCATAAATCAAGGAAATCTGTTGATGTG 480
Db 1008 GTGGCGTTTACACATATAAGAAAAAACCATAAATCAAGGAAATCTGTTGATGTG 1067
QY 481 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACAGGTCTCAA 540
Db 1068 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACAGGTCTCAA 1127
QY 541 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 600
Db 1128 GATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 1187
QY 601 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGACGTGAC 660
Db 1188 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGACGTGAC 1247
QY 661 TCCTCAATCTGCTCATGACATTAAGCAATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
Db 1248 TCCTCAATCTGCTCATGACATTAAGCAATTTTCCGTACGATTTTACCAATGGATCAAGAG 1307
QY 721 TTTACTTACCATGTCAAAATTCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
Db 1308 TTTACTTACCGTGTAAAAATTCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 1367
QY 781 AATGACAAATTAACACATCTGATCTCTGAGAAATTAATACGTCCTTAAAAAGGG 840

```

```

Db 1368 AATGACAAATAAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 1427
QY 841 GAAAGCCGCTATGATCCCTTTGATCGAGTCACTTGAAACTGTTCACCATCAATACGTT 900
Db 1428 GAAAGCCGCTATGATCCCTTTGATCGAGTCACTTGAAACTGTTCACCATCAATACGTT 1487
QY 901 GATGTCACACCAACCAAGTCTTAAAGCGAGCAGCTCTTAACAGTAGCGAACGTAAC 960
Db 1488 GATGTCACACCAACCAAGTCTTAAAGCGAGCAGCTCTTAACAGTAGCGAACGTAAC 1547
QY 961 TTAGACTTCAGAGATTTATACGATCTCTCGTGAATAAGCTTAACTCTTCAACAATCTC 1020
Db 1548 TTAGACTTCAGAGATTTATACGATCTCTCGTGAATAAGCTTAACTCTTCAACAATCTC 1607
QY 1021 GATGCTTTTGGTATTATGACATACCTTAACTGGAAGTAGAGGATATACGATGAC 1080
Db 1608 GATGCTTTTGGTATTATGACATACCTTAACTGGAAGTAGAGGATATACGATGAC 1667
QY 1081 ACCAAACGCTATCAACCGTTTATATGGCAACGACCCGAGGAGAGATGCTAGCTAT 1140
Db 1668 ACCAAACGCTATCAACCGTTTATATGGCAACGACCCGAGGAGAGATGCTAGCTAT 1727
QY 1141 CATTAGCCTATGAT 1155
Db 1728 CATTAGCCTGTTGTT 1742

RESULT 7
US-09-940-235-9
; Sequence 9, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-9

Query Match      89.3%; Score 1111.8; DB 10; Length 1541;
Best Local Similarity 97.7%; Pred. No. 4.2e-299;
Matches 1128; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATAGTTGTT 60
Db 33 ATAGCTGGTCTGCTGATGAGTCTAGATCTGCTCTTCTGTAATTAACAGCCAAATGGTTGTT 92
QY 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 120
Db 93 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAATC 152
QY 121 GACCTAACATCAGACCTGCTCATGGAGGAAAGACAGACAGCAAGCTTAAGTCCAAATCA 180

```

```
Db 153 GATCTAAATCAGCAGCTGCTCATGAGGAAACAGACAGAGCTTAAGTCCAAAATCA 212
Qy 181 AAACATTTTGTCTACTGATAGTGGCGGATGCCACATATAAACTTGAAGAAGCTGACTACTA 240
Db 213 AAACATTTTGTCTACTGATAGTGGCGGATGTGCATATAAACTTGAAGAAGCTGACTACTA 272
Qy 241 AAGGCTATTCAAGAAACAATTTGATCGCTTAACGTCACAGTAACACGACTACTTTGAGGTC 300
Db 273 AAGGCTATTCAAGAAACAATTTGATCGCTTAACGTCACAGTAACACGACTACTTTGAGGTC 332
Qy 301 ATTGATTTTGCAGGATGCAACCATTTACTGATCGAAACGGGAAGCTGCTACTTTGCTGAC 360
Db 333 ATTGATTTTGCAGGATGCAACCATTTACTGATCGAAACGGGAAGCTGCTACTTTGCTGAC 392
Qy 361 AAGATGTTTGGTAACTTTCGCGACCCACCTGTCACAGAAATTTTGTCTAAGCGGACAT 420
Db 393 AAGATGTTTGGTAACTTTCGCGACCCACCTGTCACAGAAATTTTGTCTAAGCGGACAT 452
Qy 421 GTGCGGTTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
Db 453 GTGCGGTTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 512
Qy 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 540
Db 513 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 572
Qy 541 GATATCTAAGCTATTGAACACACTAGTATCTCGGTGACACCATCATCTCAAGAAATTA 600
Db 573 GATATCTAAGCTATTGAACACACTAGTATCTCGGTGACACCATCATCTCAAGAAATTA 632
Qy 601 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTTATACGATTTATGAACGTGAC 660
Db 633 GCTCAAGCACAAGCAATTTTAAACAAAACCCAGGCTTATACGATTTATGAACGTGAC 692
Qy 661 TCTCTAATCTGCTACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
Db 693 TCTCTAATCTGCTACTCATGCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 752
Qy 721 TTTTACTTACCATCTCAAAATCGGAAACAGCTTATGAGATCAATAAAAAATCTGCTCTG 780
Db 753 TTTTACTTACCGTGTAAATTCGGAAACAGCTTATGAGATCAATAAAAAATCTGCTCTG 812
Qy 781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
Db 813 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 872
Qy 841 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACTGTTTACCATCAATAGCTT 900
Db 873 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACTGTTTACCATCAATAGCTT 932
Qy 901 GATGTCACACCAACCAATTTGCTTAAAGCGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
Db 933 GATGTCACACCAACCAATTTGCTTAAAGCGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 992
Qy 961 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTAAATCTTCAACAATCTC 1020
Db 993 TTAGACTTCAGAGATTTATACGATCTCTGATGATAGGCTAAATCTTCAACAATCTC 1052
Qy 1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGATTAATCAGATGAC 1080
Db 1053 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGATTAATCAGATGAC 1112
Qy 1081 ACCAACCGTATCATACCGTTTATGCGCAAGCGCCGAGGAGAGATGCTAGCTAT 1140
Db 1113 ACCAACCGTATCATACCGTTTATGCGCAAGCGCCGAGGAGAGATGCTAGCTAT 1172
Qy 1141 CATTTAGCCTATGAT 1155
Db 1173 CATTTAGCCTGCTG 1187
```

RESULT 8

```
US-10-369-493-29120/c
; Sequence 29120, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 29120
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-10-369-493-29120
```

```
Query Match 3.5%; Score 43.4; DB 16; Length 556;
Best Local Similarity 60.7%; Pred. No. 0.21;
Matches 71; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
Qy 688 ATTTCCGTTACGATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGAA 747
Db 292 ATTTTCAGATGTTTTTTCGATTCATCAACATTTTACCAGTCTCAAAAACAGCAA 233
Qy 748 CAAGCTTATGAGATCAATAAAAAATCTGCTGTAATGAAGAATAAACAACACTGAC 804
Db 232 AAAAATTTTAAAGTTACCAAACTATTTCCCTCAGAAAAAACTTACAAACCAATCAC 176
```

```
RESULT 9
US-10-221-613-270/c
; Sequence 270, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 270
; LENGTH: 5504
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-270
```

```
Query Match 3.4%; Score 42; DB 13; Length 5504;
Best Local Similarity 46.8%; Pred. No. 2;
Matches 132; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
```

538 AAGCATACTAAGCTATTGAAAAAAGCTAGCTATTCGGTGACACCAATCAATCTCAAGATTAA 597

1317 AAAAATATCAAACTAAACAATCTCTTTTAAATCTCTTTTCTTCTTACTAATATCAAAAAATA 1258

598 CTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCCAGGCTATACGATTTATGAACGT 657

1257 CCACCGAAACAAATAAATACTTTTAAATAAATAAATAAACAACAACTCTTATATAAAACTT 1198

658 GACTCTCAATCGTCACTCATGACAATGACATTTTCGGTACGATTTTACCAATGGATCAA 717

1197 CAAAAAATAAACAACGACCTTAAAAAATATATCTACGATCAAAATTTTAACTCTTAAAAAAA 1138

718 GAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTTATGAGATCAATAAAAATCTGGT 777

1137 AATAATCTTAAAAAACAACAATAAATAAATAAATAAATAAATAAATAAATAAAGCGT 1078

778 CTGAATGAAGAAATAAACAACACTGACCTCATCTCTGAGAAA 819

1077 AAATATAAAAAAATAAAAAAACAAGAAATAAAAACCTTAAAAA 1036

RESULT 10

US-10-311-455-340/c

; Sequence 340, Application US/10311455

; Publication No. US20030143606A1

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determination of the Cytosine Methylation of the Cytosine

; FILE REFERENCE: 5013.1014

; CURRENT APPLICATION NUMBER: US/10/311.455

; PRIOR FILING DATE: 2002-12-16

; PRIOR APPLICATION NUMBER: PCT/EP01/07537

; PRIOR FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 2424

; SEQ ID NO 340

; LENGTH: 6071

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-340

Query Match 3.4%; Score 42; DB 15; Length 6071;

Best Local Similarity 46.3%; Pred.No. 2.1;

Matches 138; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

551 TATTGAAAAACACTAGCTATTCGGTGACACCACTACATCTCAAGAATTACTAGCTCAAGCAC 610

2019 TATAAAAACTAATCTTAAAAATAAATAATCAATTTATTTCTTAAATATATAAATAATATAAATA 1960

611 AAGCATTTTAAACAAACCCACCCAGGCTATGATTTATGACGATGATCTCTCAATCG 670

1959 AAACCAACTCAATCCCAACTTAATAAAAAACGAAAAAATAATAAATACCTACCTCAACT 1900

671 TCATCTATGACATGATTCATTTCCGTACGATTTTACCAATGGATCAAGCTTTACTTACC 730

1899 AACTAATAATAATAAATAAACAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1840

731 ATGTCAAAAAATCGGGAACAAGCTTTATGAGATCAATAAAAATCTGGTCTTCAATGAAGAAA 790

1839 TAAAAAATAATCCGATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1780

791 TAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGGGAAAAAGCC 848

1779 AACCAAAACTAATAAATCCCAACTCCGAAATTAATAAATAAATAAATAAATAAATAAATAAATAA 1722

```

RESULT 11
US-10-240-485-36/c
; Sequence 36, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: Metastasis
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 36
; LENGTH: 6071
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-36

Query Match      3.4%; Score 42; DB 15; Length 6071;
Best Local Similarity 46.3%; Pred. No. 2.1;
Matches 138; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY      551 TATTGAAACACTAGCTATCGGTGACACCATCATCTCGAAGATTTACTAGCTCAAGCAC 610
Db      2019 TATAAAAAAATAAATCTTAAAAATAAAATCAATTTATTTCTTAAATATATAAAATATAAAAAA 1960

QY      611 AAAGCACTTTTAAACAAAAACCCAGGCTATACGATTTATGAACGTGACTCTCTCAATCG 670
Db      1959 AAACCAACTCAATCCAACTAATAAAAAACGAAAAAATAATAATACCTACCTCACTT 1900

QY      671 TCACCTCATGCAATGACATTTTCGGTAGAGTTTACCAATGGATCAAGAGTTTACTTACC 730
Db      1899 AAATCTAAATAAATATAAAAAAATAAAACATAAACAAAAATATAAATCTAACATCTCAACC 1840

QY      731 ATCTCAAAATCGGACACAGCTTTATGAGATCAATAAAAATCTGGTCTGATGAAGAA 790
Db      1839 TAAAAAAAATCCCGATTAAATATAAAAAAACAACAAAAAATAAATCTTAAAAAATAA 1780

QY      791 TAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGGGGAAAAGCC 848
Db      1779 AACCAAAAATAAAAATCCCAACTCCGAAATATAAAAAATATAAAACGAAAAAACC 1722

RESULT 12
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

APPLICANT: Epigenomics AG
TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC

FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match 3.4%; Score 42; DB 15; Length 3673778;
Best Local Similarity 51.2%; Pred. No. 91;
Matches 124; Conservative 0; Mismatches 115; Indels 3; Gaps 1;
QY 763 AATAAAATCTGGTCTGAATCAAGAAATAAACAACACTGACCTGATCTCTCAGAAATAT 822
Db 1605692 AATATCAAACTACTATATTTAAAAATAAABAAAACCTTACTCTTAATAAAAAATAA 1605633
QY 823 TAGCTCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGGAGTCACCTTGAAACTG 882
Db 1605632 TAAATC--AAACTAAAAAAACAATAACGATAAAAAATATTTAAATATTTTACAACC 1605576
QY 883 TTCACATCAAAATAGCTTGCATCTCAACACCAACGAATTCCTAAAAAGCGAGCTCTTTA 942
Db 1605575 CTCCTTAAAAATACATCAATTTCAAACATCTATCTACCAAAAAAACAACCACTT 1605516
QY 943 ACAGTAGCAAGCTAGTACCTTACAGCTTACAGATTTATACGATCTCTGTCGATAAGGCTAAA 1002
Db 1605515 ACGCCTTACCCCTCAATTTAAACCCCTATAATTTCTATATATCTTAAAAAATAATCAAT 1605456
QY 1003 CT 1004
Db 1605455 CT 1605454

RESULT 13
US-10-311-455-658/c
; Sequence 658, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 658
; LENGTH: 6392
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-658

Query Match 3.4%; Score 41.8; DB 15; Length 6392;
Best Local Similarity 50.2%; Pred. No. 2.5;
Matches 103; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 546 TAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGATTACTAGCTCA 605
Db 377 TAACCTACTAAAAACCCCTTCTTAACGAACCTAACTCTTCTTAAACTAAATCTTAA 318
QY 606 AGCACAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACGTCCTCTC 665
Db 317 ACTAACACACTAATAACTCTTCCCGATCAATTAATTAACAAACTTTACCCATA 258
QY 666 AATCGTCACTGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGGTTTAC 725
Db 197 TTATCATAAAAAATCTTAATCTA 173

Db 257 AATAATTTCAACAGAAATATCAAAATCTTCCGATATCTTCCGATATCTTCAAAATAAATC 198
QY 726 TTACCATGTCAAAAATCGGGAACAA 750
Db 197 TTATCATAAAAAATCTTAATCTA 173

RESULT 14
US-10-240-485-60/c
; Sequence 60, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 60
; LENGTH: 6392
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-60

Query Match 3.4%; Score 41.8; DB 15; Length 6392;
Best Local Similarity 50.2%; Pred. No. 2.5;
Matches 103; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 546 TAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGATTACTAGCTCA 605
Db 377 TAACCTACTAAAAACCCCTTCTTAACGAACCTAACTCTTCTTAAACTAAATCTTAA 318
QY 606 AGCACAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACGTCCTCTC 665
Db 317 ACTAACACACTAATAACTCTTCCCGATCAATTAATTAACAAACTTTACCCATA 258
QY 666 AATCGTCACTGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGGTTTAC 725
Db 257 AATAATTTCAACAGAAATATCAAAATCTTCCGATATCTTCCGATATCTTCAAAATAAATC 198
QY 726 TTACCATGTCAAAAATCGGGAACAA 750
Db 197 TTATCATAAAAAATCTTAATCTA 173

RESULT 15
US-10-311-455-61/c
; Sequence 61, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 5, 2004, 23:29:56 ; Search time 552 Seconds
(without alignments)
9581.534 Million cell updates/sec

Title: US-09-882-509-3

Perfect score: 1245

Sequence: 1 attgttgacgtgagtggtct.....ataaccctaacgacaaataa 1245

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1241.8	99.7	1245	2	AAQ20665 SKC-2 str
2	1241.8	99.7	1335	2	AAQ12156 Streptoki
3	1241.8	99.7	1458	2	AAQ12162 Factor Xa
4	1241.8	99.7	1512	2	AAQ12158 Streptoki
5	1241.8	99.7	2589	2	AAQ12160 OmpAL str
6	1238.8	99.5	1467	2	AAQ12490 Factor Xa
7	1224.2	98.3	1245	3	AAQ37633 S. equis
8	1224.2	98.3	1473	3	AAQ05603 Streptoki
9	1222.6	98.2	1377	3	AAQ37622 Streptoki
10	1222.6	98.2	1782	3	AAQ37642 Chimeric
11	1222.6	98.2	2030	2	AAQ11651 PB-PB-SK
12	1221.2	98.1	1242	2	AAQ80492 Streptoco
13	1221.2	98.1	1254	6	ABA05546 Streptoki
14	1221.2	98.1	2385	2	AAQ80497 Streptoki
15	1221.2	98.1	8893	6	ABA05547 Maxadilan
16	1221.6	98.0	1242	2	AAQ16632 Streptoco
17	1219.4	97.9	1323	2	AAQ29961 Vector pS
18	1216.4	97.7	1242	2	AAQ16633 Streptoco
19	1210.2	97.2	2566	3	AAQ77778 Coding se
20	1209.8	97.2	1327	3	AAQ37628 Streptoki
21	1206.4	96.9	1209	3	AAQ299249 DNA encod
22	1203.8	96.7	1245	3	AAQ299250 DNA encod
23	1192.2	95.8	1407	1	AAQ70106 DNA encod

24	1182.6	95.0	2568	1	AAQ50493 Sequence
25	1170.4	94.0	1242	5	AAF82144 Mutant st
26	1127.8	90.6	1661	3	AAQ37637 Chimeric
27	1124.6	90.3	2096	3	AAQ37643 Chimeric
28	1116	89.6	1122	3	AAQ299251 DNA encod
29	1116	89.6	1158	3	AAQ299252 DNA encod
30	1111.8	89.3	1541	3	AAQ37644 Chimeric
31	1103.4	88.6	1119	2	AAQ12159 Truncated
32	1103.2	88.6	2253	2	AAQ12161 Met-core
33	1094.6	87.9	1473	2	AAQ05604 Streptoki
34	1093.2	87.8	1320	6	ABN70192 Streptoco
35	1050.4	84.4	1068	2	AAQ80493 Recombina
36	1045.8	84.0	2208	2	AAQ83589 Recombina
37	811.4	65.2	1262	2	AAQ10230 Synthetic
38	440.4	35.4	450	2	AAQ80494 Recombina
39	46.2	3.7	2000	7	ADA71938 Rice gene
40	43.8	3.5	2000	7	ADA71938 Rice gene
41	42	3.4	1180	2	AAQ84163 Streptoco
42	42	3.4	1181	2	AAQ84164 Streptoco
43	42	3.4	5504	6	ABL70350 Chemical
44	42	3.4	5504	6	AAQ61309 Human imm
45	42	3.4	6071	6	ABL32367 Human imm

ALIGNMENTS

RESULT 1

AAQ20665

ID AAQ20665 standard; DNA; 1245 BP.

AC AAQ20665;

DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 21-APR-1992 (first entry)

XX SKC-2 streptokinase gene from S.equisimilis.

XX plasminogen activator; coronary thrombosis; ATCC-9542; ss.

OS Streptococcus dysgalactiae subsp. equisimilis.

PN AU9178101-A.

PD 28-NOV-1991.

PF 31-MAY-1991; 91AU-00078101.

PR 23-MAY-1990; 90CU-00000090.

XX (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.

DR WPI; 1992-024716/04.

XX P-PSDB; AAR20202.

PT Streptokinase C-2 gene from S equisimilis type C - plasmids and transformants for large scale intra- and extracellular expression of SKC-2 useful in thrombolytic agents.

PS Claim 6; Page 14; 28pp; English.

CC The SKC-2 gene is isolated from S.equisimilis type C by gene amplification from the synthetic primers SK1, SK2 and SK3 (see AAQ20666-8). The SKC-2 gene was inserted into yeast expression vector pPS-7 for extracellular expression of streptokinase and into pNAO for intracellular expression in yeast. For expression in bacteria, the SKC-2 gene was inserted in plasmid pEK33, between a trp promoter and a T4 terminator. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)

XX Sequence 1245 BP; 428 A; 273 C; 232 G; 312 T; 0 U; 0 Other;

SQ

Query Match 99.7%; Score 1241.8; DB 2; Length 1245;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATTGCTGAGCTGAGTGGCTGCTAGACCGTCTGTCACACGCTCAACACGCAATAGTTGTT 60
Db 1 ATTGCTGAGCTGAGTGGCTGCTAGACCGTCTGTCACACGCTCAACACGCAATAGTTGTT 60
QY 61 AGCGTTGCTGCTACTGTTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATTT 120
Db 61 AGCGTTGCTGCTACTGTTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAATTT 120
QY 121 GACCTAACATCAGACGCTGCTCTGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 180
Db 121 GACCTAACATCAGACGCTGCTCTGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 180
QY 181 AAACCATTTGCTACTGATAGTGGCGCGATGCCATAACTTGAAGAAAGCTGACTTACTA 240
Db 181 AAACCATTTGCTACTGATAGTGGCGCGATGCCATAACTTGAAGAAAGCTGACTTACTA 240
QY 241 AAGGCTATTCAAGAACCAATTGATCGCTAACTGCTCAAGTAAACGACGACTACTTTGAGGTC 300
Db 241 AAGGCTATTCAAGAACCAATTGATCGCTAACTGCTCAAGTAAACGACGACTACTTTGAGGTC 300
QY 301 ATTGATTTGCAAGCGATGCAACCAATTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 360
Db 301 ATTGATTTGCAAGCGATGCAACCAATTACTGATCGAAACGCAAGGCTTACTTTGCTGAC 360
QY 361 AAGATAGTTGCGTAACTGCTGAGCCCAAGCTGCTCAAGAAATTTTGTCAAGCGGACAT 420
Db 361 AAGATAGTTGCGTAACTGCTGAGCCCAAGCTGCTCAAGAAATTTTGTCAAGCGGACAT 420
QY 421 GTGCGGCTAGACCATATAAAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
Db 421 GTGCGGCTAGACCATATAAAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
QY 481 GAATATAGTACGATTTACTCCCTTAAACCTGATGACGATTTCAAGCAGGCTCTCAA 540
Db 481 GAATATAGTACGATTTACTCCCTTAAACCTGATGACGATTTCAAGCAGGCTCTCAA 540
QY 541 GATACTAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA 600
Db 541 GATACTAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTACTA 600
QY 601 GCTCAAGCACAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACGTCGAC 660
Db 601 GCTCAAGCACAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACGTCGAC 660
QY 661 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 720
Db 661 TCCTCAATCGTCACTGATGACATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 720
QY 721 TTTTACTTACCATGTCAAAAATCGGAAACAGCTTATGAGATCAATAAATAATCTGGTCTG 780
Db 721 TTTTACTTACCATGTCAAAAATCGGAAACAGCTTATGAGATCAATAAATAATCTGGTCTG 780
QY 781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTAGTCTTAAAAAGGG 840
Db 781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTAGTCTTAAAAAGGG 840
QY 841 GAAAAGCGGTATGATCCCTTTGATCGCATGCTACTTGAACCTGTTCCACATCAATAGCTT 900
Db 841 GAAAAGCGGTATGATCCCTTTGATCGCATGCTACTTGAACCTGTTCCACATCAATAGCTT 900
QY 901 GATGTCAACCAAGCAATTTGCTTAAAAAGCGAGCAGCTCTTAAACAGCTAGGCAACGTAAC 960
Db 901 GATGTCAACCAAGCAATTTGCTTAAAAAGCGAGCAGCTCTTAAACAGCTAGGCAACGTAAC 960
QY 961 TTAGACTTCAGAGATTTATAGATCTCTGATGAAGGCTAACTACTTACAAATCTC 1020
Db 961 TTAGACTTCAGAGATTTATAGATCTCTGATGAAGGCTAACTACTTACAAATCTC 1020
QY 1021 GATGCTTTTGGTATTATCGACTATACCTTAACTGGAAGAGTAGAGATAATCACGATGAC 1080
```

```
Db 1021 GATGCTTTTGGTATTATCGACTATACCTTAACTGGAAGAGTAGAGATAATCACGATGAC 1080
QY 1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1140
Db 1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGAGAGAATGCTAGCTAT 1140
QY 1141 CATTAGCCCTATGATAAAGATCGTTTATACCGAAGAAAGAGAGAGTTTACAGCTACCTG 1200
Db 1141 CATTAGCCCTATGATAAAGATCGTTTATACCGAAGAAAGAGAGAGTTTACAGCTACCTG 1200
QY 1201 CGTTATACAGGACACCTATACCTGATAACCTTAAACGACAATAA 1245
Db 1201 CGTTATACAGGACACCTATACCTGATAACCTTAAACGACAATAA 1245

RESULT 2
AAQ12156
ID AAQ12156 standard; DNA; 1335 BP.
XX
XX AAQ12156;
XX AC
XX 24-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 17-SEP-1991 (first entry)
XX
XX Streptokinase gene.
XX
XX Fusion protein; blood clotting; coagulation; fibrinolysis;
XX antithrombotic; thrombolysis; streptokinase; ss.
XX
XX Streptococcus dysgalactiae subsp. equisimilis; ATCC 9542 or ATCC 10009.
XX
XX Key Location/Qualifiers
FH CDS 7..1329 /*tag= a
FT sig_peptide 7..84 /*tag= b
FT mat_peptide 85..1326 /*tag= c
FT
XX
XX WO9109125-A.
XX
XX 27-JUN-1991.
XX
XX 07-DEC-1989; 89GB-00027722.
XX
XX 07-DEC-1989; 89GB-00027722.
XX
XX 07-DEC-1990; 90WO-GB001911.
XX
XX (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
XX
XX Dawson KM, Hunter MG, Czapliewsk LG;
XX
XX WPI; 1991-208151/28.
XX
XX P-PSDB; AAR12889.
XX
XX Fusion protein cleavage by blood clotting enzyme - for prodn. of
XX fractions having greater antithrombotic activity for therapy and
XX prophylaxis.
XX
XX Disclosure; Page 80; 115pp; English.
XX
XX The sequence was obt'd. from PCR amplified chromosomal DNA from S.
XX equisimilis (Lancefield's Gp C ) ATCC 10009 or ATCC 9642. The primers
XX used for the PCR were based on the published DNA sequence of S.
XX equisimilis strain H46A (Malke, H., Roe, B., and Ferretti, J.J., Gene 34
XX 357-362 [1985]). See Swiss-Prot K02986 and Geneseq N70106. The sequence
XX is an EcoRI-BamHI fragment. The gene can be used to construct expression
XX vectors in which the streptokinase gene is linked to a second gene
XX encoding e.g. another streptokinase protein, hirudin, or a streptokinase-
XX like protein, via a linking sequence encoding a cleavage site for e.g.
XX factor Xa or thrombin. The enzymes which cleave the fusion protein are
```

CC present at the site of the target thrombus so the active agents are
CC released specific- ally at the place where clot formation is occurring.
CC See also Q12153-Q12155, Q12158-Q12162 and Q12490. (Updated on 25-MAR-2003
CC to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 1335 BP; 453 A; 287 C; 250 G; 345 T; 0 U; 0 Other;
Query Match 99.7%; Score 1241.8; DB 2; Length 1335;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCTCAACACGCAATAGTTGTT 60
Db |||
85 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCTCAACACGCAATAGTTGTT 144
QY 61 AGCGTTGCTGGTACTGTGTGGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 120
Db |||
145 AGCGTTGCTGGTACTGTGTGGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 204
QY 121 GACCTAACATCAGACCTGCTCATGGAGGAAAGACAGACGCAAGCTTAACTCCAAATCA 180
Db |||
205 GACCTAACATCAGACCTGCTCATGGAGGAAAGACAGACGCAAGCTTAACTCCAAATCA 264
QY 181 AATCCATTTGCTACTGATAGTGGCGGATGCCACATAAATTCGAAAAGCTGACTTACTA 240
Db |||
265 AATCCATTTGCTACTGATAGTGGCGGATGCCACATAAATTCGAAAAGCTGACTTACTA 324
QY 241 AAGGCTATTCAAGACAAATTTGATCGCTTAACGTCACAGTAACGACGACTTCTTGGTTC 300
Db |||
325 AAGGCTATTCAAGACAAATTTGATCGCTTAACGTCACAGTAACGACGACTTCTTGGTTC 384
QY 301 ATTGATTTTGCAGCGATGCAACATTTACTGTGATCGAAACGGCAAGTCTACTTTGCTGAC 360
Db |||
385 ATTGATTTTGCAGCGATGCAACATTTACTGTGATCGAAACGGCAAGTCTACTTTGCTGAC 444
QY 361 AAGATGTTTGGTAACTTTGCGGACCGAACCTGTGTCAGAAATTTTGTCTAAAGCGACAT 420
Db |||
445 AAGATGTTTGGTAACTTTGCGGACCGAACCTGTGTCAGAAATTTTGTCTAAAGCGACAT 504
QY 421 GTGCGGTTAGACCATATAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTC 480
Db |||
505 GTGCGGTTAGACCATATAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTC 564
QY 481 GAATATCTGTACAGTTTACTTCCCTTAAACCCCTGTGATGACGATTTTCAGACGAGTCTCAA 540
Db |||
565 GAATATCTGTACAGTTTACTTCCCTTAAACCCCTGTGATGACGATTTTCAGACGAGTCTCAA 624
QY 541 GATACTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 600
Db |||
625 GATACTAAGCTATTGAACACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA 684
QY 601 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTTATGAACGTGAC 660
Db |||
685 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTTATGAACGTGAC 744
QY 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTCAGATTTTACCAATGGATCAAGAG 720
Db |||
745 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTCAGATTTTACCAATGGATCAAGAG 804
QY 721 TTTTACTTACCATGTCAAAATCGGGACAAAGCTTATGAGATCAATATAAAATCTGTCCTG 780
Db |||
805 TTTTACTTACCATGTCAAAATCGGGACAAAGCTTATGAGATCAATATAAAATCTGTCCTG 864
QY 781 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
Db |||
865 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 924
QY 841 GAAAACCGGTATGATCCCTTTTGTATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 900
Db |||
925 GAAAACCGGTATGATCCCTTTTGTATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 984
QY 901 GATGTCACACCAACGAATTTGCTTAAAGCGGACGACTCTTAAACGCTAGCGAACCTAAC 960
Db |||

Db 985 GATGTCACACCAACGAATTTGCTTAAAGACGAGCAGCTCTTAAACAGCTAGCGACGTAAC 1044
QY 961 TTAGACTTCAGAGATTTTACGATCTCTGTGATAAGGCTAAACTACTCTTACAAATCTC 1020
Db |||
1045 TTAGACTTCAGAGATTTTACGATCTCTGTGATAAGGCTAAACTACTCTTACAAATCTC 1104
QY 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAGGATAATCAGATGAC 1080
Db |||
1105 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAAGTAGAGGATAATCAGATGAC 1164
QY 1081 ACCAACCGTATATACACCGTTTATATGGCAAGCACCAGGAGAGAGATGCTAGCTAT 1140
Db |||
1165 ACCAACCGTATATACACCGTTTATATGGCAAGCACCAGGAGAGAGATGCTAGCTAT 1224
QY 1141 CATTTAGCCCTATGATAAAGATCGTTTATACCGAAGAAGACGAGAGATTTACAGCTACCTG 1200
Db |||
1225 CATTTAGCCCTATGATAAAGATCGTTTATACCGAAGAAGACGAGAGATTTACAGCTACCTG 1284
QY 1201 CGTTATACAGGGACACCTATACCTGATTAACCTTAAACGACAAATAA 1245
Db |||
1285 CGTTATACAGGGACACCTATACCTGATTAACCTTAAACGACAAATAA 1329
RESULT 3
AAQ12162
ID AAQ12162 standard; DNA; 1458 BP.
XX
AC AAQ12162;
XX
DT 25-MAR-2003 (revised)
DT 17-SEP-1991 (first entry)
XX
Factor Xa-cleavable hirudin-IEGR-streptokinase gene.
XX
Fusion protein; blood clotting; coagulation; fibrinolysis;
KW antithrombotic; thrombolysis; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1452
FT /tag= a
FT mat_peptide 1..195
FT /tag= b
FT /label= hirudin HV-1
FT misc_RNA 196..207
FT /tag= c
FT /label= linker
FT /note= "encodes factor Xa cleavage site"
FT mat_peptide 208..1449
FT /tag= d
FT /label= streptokinase
XX
PN WO9109125-A.
XX
PD 27-JUN-1991.
XX
PP 07-DEC-1989; 89GB-00027722.
XX
PR 07-DEC-1989; 89GB-00027722.
PR 07-DEC-1990; 90WO-GH001911.
XX
PA (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
XX
PI Dawson KM, Hunter MG, Czaplewsk LG;
XX
DR WPI; 1991-208151/28.
XX
P-PSDB; AAR12885.
PT Fusion protein cleavage by blood clotting enzyme - for prodn. of
PT fractions having greater antithrombotic activity for therapy and
PT prophylaxis.
XX

PS Disclosure; Page 96; 115pp; English.

XX The sequence of the synthetic hirudin HV-1 genes was designed based on
CC the published amino acid sequence (Dodd J., et al PEBB Letters 165 180
CC (1984)). The sequence of streptokinase was obtained from PCR amplified
CC chromosomal DNA from *S. equisimilis* ATCC 10009 or ATCC 9642. The primers
CC used for the PCR were based on the published DNA sequence of *S.*
CC *equisimilis* strain H46A (Walke, H., Roe, B., and Ferretti, J.J., Gene 34
CC 357-362 [1985]). The two sequences were used to construct an expression
CC vector in which the hirudin gene is linked to the streptokinase gene via
CC a linking sequence encoding a cleavage site for factor Xa. The factor Xa
CC is present at the site of the target thrombus so the active agents are
CC released specifically at the place where clot formation is occurring. See
CC also AAQ12153-Q12156, AAQ12158-Q12161 and AAQ12490. (Updated on 25-MAR-
CC 2003 to correct PA field.)
XX

SQ Sequence 1458 BP; 491 A; 316 C; 290 G; 361 T; 0 U; 0 Other;

Query Match 99.7%; Score 1241.8; DB 2; Length 1458;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATTCGTGACCTGAGTCTGCTAGACCGCTCATCTGTCAACACAGCCCAATTAGTTGTT 60
DB 208 ATTCGTGACCTGAGTCTGCTAGACCGCTCATCTGTCAACACAGCCCAATTAGTTGTT 267
QY 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 120
DB 268 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 327
QY 121 GACCTAATCAGACCTGCTCATGAGGAAAGACAGAGCGTTAAGTCCAAATCA 180
DB 328 GACCTAATCAGACCTGCTCATGAGGAAAGACAGAGCGTTAAGTCCAAATCA 387
QY 181 AAACCATTTGCTACTGATAGTGGCGCATGCCATCAATAACTTGAAGAGCTGACTACTA 240
DB 388 AAACCATTTGCTACTGATAGTGGCGCATGCCATCAATAACTTGAAGAGCTGACTACTA 447
QY 241 AAGCTATTCAAGAACCAATTGATCGCTAACTCAAGTACAGCAGCTACTTTGAGTC 300
DB 448 AAGCTATTCAAGAACCAATTGATCGCTAACTCAAGTACAGCAGCTACTTTGAGTC 507
QY 301 ATTGATTTGCAAGGATGCAACCAATTACTGATGAAACGCAAGTCTACTTTGCTGAC 360
DB 508 ATTGATTTGCAAGGATGCAACCAATTACTGATGAAACGCAAGTCTACTTTGCTGAC 567
QY 361 AAAGATGGTTCGGTAACTTCGCCACCACTGCTCCAAAGAAATTTTGTAAAGCGGACAT 420
DB 568 AAAGATGGTTCGGTAACTTCGCCACCACTGCTCCAAAGAAATTTTGTAAAGCGGACAT 627
QY 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
DB 628 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 687
QY 481 GAATATAGCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCAGGTTCTCAA 540
DB 688 GAATATAGCTGACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCAGGTTCTCAA 747
QY 541 GATATAGCTGATTTGAAACCACTAGCTATCGGTGACACCATCAATCTCAAGAAATTACTA 600
DB 748 GATATAGCTGATTTGAAACCACTAGCTATCGGTGACACCATCAATCTCAAGAAATTACTA 807
QY 601 GCTCAAGCAAAAGCATTTTAAACAAACCCAGCCTATACGATTTATGAACGTGAC 660
DB 808 GCTCAAGCAAAAGCATTTTAAACAAACCCAGCCTATACGATTTATGAACGTGAC 867
QY 661 TCCTCAATCGTCATCATGCAATGACATTTTCCGTAGATTTTACCAATCGATCAAGAG 720
DB 868 TCCTCAATCGTCATCATGCAATGACATTTTCCGTAGATTTTACCAATCGATCAAGAG 927
QY 721 TTTTACTTACCATGTCAAAATTCGGGAAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
DB 928 TTTTACTTACCATGTCAAAATTCGGGAAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 987

QY 781 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 840
DB 988 AATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 1047
QY 841 GAAAGCCGTTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTCAACATCAATAGCTT 900
DB 1048 GAAAGCCGTTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTCAACATCAATAGCTT 1107
QY 901 GATGTCAACACCAACGAATTGCTTAAAGGAGGAGCAGCTCTTAACAGCTAGGAGACGTAAC 960
DB 1108 GATGTCAACACCAACGAATTGCTTAAAGGAGGAGCAGCTCTTAACAGCTAGGAGACGTAAC 1167
QY 961 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGCTAAACTACTCTACAACAATCTC 1020
DB 1168 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGCTAAACTACTCTACAACAATCTC 1227
QY 1021 GATGTTTTGTTATATGAGCTATACCTTAACCTGAAAAAGTAGAGGATAATACAGATGAC 1080
DB 1228 GATGTTTTGTTATATGAGCTATACCTTAACCTGAAAAAGTAGAGGATAATACAGATGAC 1287
QY 1081 ACCAACCGTATCAACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
DB 1288 ACCAACCGTATCAACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1347
QY 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAACGAGAACTTTACAGCTACCTG 1200
DB 1348 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAACGAGAACTTTACAGCTACCTG 1407
QY 1201 CGTTATACAGGACCACTATACCTGATTAACCTTAACGACAAATAA 1245
DB 1408 CGTTATACAGGACCACTATACCTGATTAACCTTAACGACAAATAA 1452

RESULT 4

AAQ12158
ID AAQ12158 standard; DNA; 1512 BP.

XX AAQ12158;

XX 24-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 17-SEP-1991 (first entry)

Streptokinase gene fused to yeast alpha factor secretion sequence.

Fusion protein; blood clotting; coagulation; fibrinolysis;
antithrombotic; thrombolysis; streptokinase; ss.

Streptococcus dysgalactiae subsp. *equisimilis*; ATCC 9542 or ATCC 10009.

Key Location/Qualifiers

CDS 7..1506 /*tag= a
sig_peptide 7..261 /*tag= b
mat_peptide 262..1503 /*tag= c
/label= mature streptokinase

W09109125-A.

27-JUN-1991.

07-DEC-1989; 89GB-00027722.

07-DEC-1989; 89GB-00027722.

07-DEC-1990; 90WO-GB001911.

(BRBI-) BRITISH BIO-TECHNOLOGY LTD.

Dawson KM, Hunter MG, Czaplinski LG;

XX WPI: 1991-208151/28.
 DR P-PSDB; AAR12891.
 XX
 PT Fusion protein cleavage by blood clotting enzyme - for prodn. of
 PT fractions having greater antithrombotic activity for therapy and
 PT prophylaxis.
 XX
 PS Disclosure; Page 86; 115pp; English.
 XX
 CC The streptokinase sequence was obt. from PCR amplified chromosomal DNA
 CC from *S. equisimilis* (lancefield's Gp C) ATCC 10009 or ATCC 9642 (the
 CC primers used for the PCR were based on the published DNA sequence of *S.*
 CC *equisimilis* strain H46A (Walke, H., Roe, B., and Ferretti, J.J., Gene 34
 CC 357-362 (1985)). The gene was fused to DNA encoding the yeast alpha
 CC factor pre-pro-secretion sequence in an expression vector, pSMD1/152, for
 CC prodn. of recombinant strepto- kinase in *S. cerevisiae* strain Buz168. See
 CC also AAQ12153-Q12156, AAQ12159-Q12162 and AAQ12490. (Updated on 25-MAR-
 CC 2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
 CC field)
 XX
 SQ Sequence 1512 BP; 508 A; 323 C; 288 G; 393 T; 0 U; 0 Other;
 Query Match 99.7%; Score 1241.8; DB 2; Length 1512;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGACCAATTAGTTGT 60
 DB 262 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGACCAATTAGTTGT 321
 QY 61 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAAT 120
 DB 322 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAAT 381
 QY 121 GACCTAACATCGACCTGCTCATGGAGGAAGACAGACAGAGGCTTAAGTCAAAATCA 180
 DB 382 GACCTAACATCGACCTGCTCATGGAGGAAGACAGACAGAGGCTTAAGTCAAAATCA 441
 QY 181 AAACCATTTGCTACTGATAGTGGCGGATGCCATATACATTAACCTTGAAAGCTGACTACTA 240
 DB 442 AAACCATTTGCTACTGATAGTGGCGGATGCCATATACATTAACCTTGAAAGCTGACTACTA 501
 QY 241 AAGGCTATTCAAGAACAAATTGATCGTAAACGTCACAGTAAACGACGACTACTTTGAGTGC 300
 DB 502 AAGGCTATTCAAGAACAAATTGATCGTAAACGTCACAGTAAACGACGACTACTTTGAGTGC 561
 QY 301 ATTGATTTGCAAGGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 360
 DB 562 ATTGATTTGCAAGGATGCAACCATTTACTGATCGAAACGGCAAGTCTACTTTGCTGAC 621
 QY 361 AAAGATGTTGCGTAACCTTGGCGGCCCAACCTGTCCCAAGAAATTTTGTAAAGCGGACAT 420
 DB 622 AAAGATGTTGCGTAACCTTGGCGGCCCAACCTGTCCCAAGAAATTTTGTAAAGCGGACAT 681
 QY 421 GTGCGGTTAGACCAATATAAGAAACCAATACAAATCAAGCGAAATCTGTGTATGTG 480
 DB 682 GTGCGGTTAGACCAATATAAGAAACCAATACAAATCAAGCGAAATCTGTGTATGTG 741
 QY 481 GAATATAGTACAGTTTACTCGCTTAAACCTGTAGCGATTCAGACAGTCTCAAA 540
 DB 742 GAATATAGTACAGTTTACTCGCTTAAACCTGTAGCGATTCAGACAGTCTCAAA 801
 QY 541 GATCTAAGCTATTGAAACACTAGTCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 600
 DB 802 GATCTAAGCTATTGAAACACTAGTCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 861
 QY 601 GCTCAAGCAAAAGCAATTTTAAACAAACCCACCCAGGCTATACGATTTATGAACGTGAC 660
 DB 862 GCTCAAGCAAAAGCAATTTTAAACAAACCCACCCAGGCTATACGATTTATGAACGTGAC 921
 QY 661 TCGTCAATCGTCACTCATGACAAATGATTTTCGGTACGATTTTACCAATGGATCAAGAG 720

DB 922 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTTACGATTTTACCAATGGATCAAGAG 981
 QY 721 TTTACTTACCATGTCAAAATCGGAAACAAAGCTTATGAGATCAATAAAAATCTGTGCTG 780
 DB 982 TTTACTTACCATGTCAAAATCGGAAACAAAGCTTATGAGATCAATAAAAATCTGTGCTG 1041
 QY 781 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 840
 DB 1042 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1101
 QY 841 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCAATCAATACGTT 900
 DB 1102 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCAATCAATACGTT 1161
 QY 901 GATGTCACACCAACGAATTTGCTTAAAGCGAGCGAGCTCTTAACAGCTAGCGAACGTAAC 960
 DB 1162 GATGTCACACCAACGAATTTGCTTAAAGCGAGCGAGCTCTTAACAGCTAGCGAACGTAAC 1221
 QY 961 TTAGACTTCAGAGATTTATACGATCTCTGATAGAGGCTTAAACTACTCTACCAATCTC 1020
 DB 1222 TTAGACTTCAGAGATTTATACGATCTCTGATAGAGGCTTAAACTACTCTACCAATCTC 1281
 QY 1021 GATGCTTTGGTATTATGGACTATATACCTTTAACTGGAAGAGTAGAGGATTAATCAGATGAC 1080
 DB 1282 GATGCTTTGGTATTATGGACTATATACCTTTAACTGGAAGAGTAGAGGATTAATCAGATGAC 1341
 QY 1081 ACCAACCGTATCATACCGCTTTATATGGCAAGCGACCCGAAAGGAGAGATCTAGCTAT 1140
 DB 1342 ACCAACCGTATCATACCGCTTTATATGGCAAGCGACCCGAAAGGAGAGATCTAGCTAT 1401
 QY 1141 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAACGAGAGTTTTACAGCTACCTG 1200
 DB 1402 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAACGAGAGTTTTACAGCTACCTG 1461
 QY 1201 CGTTATACGGGACACCTATACCTGATTAACCCCTTAACGACAAATAA 1245
 DB 1462 CGTTATACGGGACACCTATACCTGATTAACCCCTTAACGACAAATAA 1506
 RESULT 5
 ID AAQ12160 standard; DNA; 2589 BP.
 XX AAQ12160;
 XX AC
 XX 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 17-SEP-1991 (first entry)
 XX
 DE OmpAL streptokinase-streptokinase gene.
 XX
 KW Fusion protein; blood clotting; coagulation; fibrinolysis;
 KW antithrombotic; thrombolysis; major outer membrane protein A; thrombin;
 KW 88.
 XX
 OS *Streptococcus dysgalactiae* subsp. *equisimilis*; ATCC 9542 or ATCC 10009.
 XX
 FH Key Location/Qualifiers
 FT CDS 4..2583
 FT sig_peptide /*tag= a
 FT 4..66
 FT /*tag= b
 FT /label= ompAL signal sequence
 FT /note= "E.coli"
 FT 67..1308
 FT mat_peptide /*tag= c
 FT /label= mature streptokinase
 FT 1339..2580
 FT /*tag= e
 FT /label= mature streptokinase
 FT 1809..1338
 FT CDS /*tag= d
 FT /label= linker

/note= "encodes thrombin cleavage site"

XX PF W09109125-A.
 XX PD 27-JUN-1991.
 XX XX 07-DEC-1989; 89GB-00027722.
 XX XX 07-DEC-1989; 89GB-00027722.
 XX PR 07-DEC-1990; 90WO-GB001911.
 XX XX (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
 XX PI Dawson KM, Hunter MG, Czaplewsk LG;
 XX WPI; 1991-208151/28.
 DR P-PSDB; AAR12893.
 XX XX Fusion protein cleavage by blood clotting enzyme - for prodn. of
 PT fractions having greater antithrombotic activity for therapy and
 PT prophylaxis.
 XX PS Disclosure; Page 90; 115pp; English.
 CC The streptokinase sequence was obtd. from PCR amplified chromosomal DNA
 CC from S. equisimilis ATCC 10009 or ATCC 9642. The primers used for the PCR
 CC were based on the published DNA sequence of S. equisimilis strain H46A
 CC (Walke, H., Roe, B., and Ferretti, J.J., Gene 34 357-362 [1985]). The
 CC gene was used to construct an expression vector in which the two
 CC streptokinase genes are linked via a link- ing sequence. This encodes a
 CC cleavage site for thrombin which, when cleaved, releases the individual
 CC proteins which have anti- thrombotic activity. The thrombin is present at
 CC the site of the target thrombus so the active agents are released
 CC specifically at the place where clot formation is occurring. See also
 CC AAQ12153012156, AAQ12158-012162 and AAQ12490. (Updated on 25-MAR-2003 to
 CC correct FA field.) (Updated on 24-OCT-2003 to standardise OS field)
 XX SQ Sequence 2589 BP; 879 A; 564 C; 500 G; 646 T; 0 U; 0 Other;
 Query Match 99.7%; Score 1241.8; DB 2; Length 2589;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 ATTCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCGCAATAGTTGTT 60
 Db 1339 ATTCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCGCAATAGTTGTT 1398
 Qy 61 AGCGTTGCTGCTACTGTTGAGGGGACGNAATCAAGACATTAGTCTTAAATTTTGAATTT 120
 Db 1399 AGCGTTGCTGCTACTGTTGAGGGGACGNAATCAAGACATTAGTCTTAAATTTTGAATTT 1458
 Qy 121 GACCTAACATCAGCACTGCTCATGGAGGAAGACAGACGAGCGCTTAAAGTCCAAATCA 180
 Db 1459 GACCTAACATCAGCACTGCTCATGGAGGAAGACAGACGAGCGCTTAAAGTCCAAATCA 1518
 Qy 181 AAACCATTTGCTACTGATAGTGGCGCGATGCCAATTAACCTTAAAGTCCAAATCA 240
 Db 1519 AAACCATTTGCTACTGATAGTGGCGCGATGCCAATTAACCTTAAAGTCCAAATCA 1578
 Qy 241 AAGGCTATTCAAGAACCAATTGATCGTAAACCTGACACGACGACGACTCTTTGAGTGC 300
 Db 1579 AAGGCTATTCAAGAACCAATTGATCGTAAACCTGACACGACGACGACTCTTTGAGTGC 1638
 Qy 301 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 360
 Db 1639 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 1698
 Qy 361 AAAGATGTTGCTGTAACCTTGGCGACCCAACTGTCGCAAGAAATTTTTCGTAAGCGGACAT 420
 Db 1699 AAAGATGTTGCTGTAACCTTGGCGACCCAACTGTCGCAAGAAATTTTTCGTAAGCGGACAT 1758
 Qy 421 GTGCGCGTTAGACCATATAAGAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480

Db 1759 GTGCGCGTTAGACCATATAAGAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 1818
 Qy 481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAACAGCGTCTCAAA 540
 Db 1819 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAACAGCGTCTCAAA 1878
 Qy 541 GATACTAAGCTATTGAAAACACTAGCTTATCGGTGACACCATCATCTCAAGAAATTAATA 600
 Db 1879 GATACTAAGCTATTGAAAACACTAGCTTATCGGTGACACCATCATCTCAAGAAATTAATA 1938
 Qy 601 GCTCAAGCACAAGCAATTTTAAACCAACCCACCCAGGCTATACGATTTATGACAGTGAC 660
 Db 1939 GCTCAAGCACAAGCAATTTTAAACCAACCCATCCAGGCTATACGATTTATGACAGTGAC 1998
 Qy 661 TCCTCAATCGTCACTCATGACAAATGACATTTCCGTACGATTTTACCAATGGATCAAGAG 720
 Db 1999 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 2058
 Qy 721 TTTACTTACCATGTCAAAAATCGGAAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780
 Db 2059 TTTACTTACCATGTCAAAAATCGGAAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 2118
 Qy 781 AATGAGAAATTAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 840
 Db 2119 AATGAGAAATTAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAGGG 2178
 Qy 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACATCAATACGTT 900
 Db 2179 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTCAACATCAATACGTT 2238
 Qy 901 GATGTCAACACCAACGAATTTGCTTAAAGCGAGGAGCTCTTAAACAGCTAGCGAACGTAAC 960
 Db 2239 GATGTCAACACCAACGAATTTGCTTAAAGCGAGGAGCTCTTAAACAGCTAGCGAACGTAAC 2298
 Qy 961 TTAGACTTTCAGAGATTTATACGATCCCTCGTGATAAGCTAACTACTCTCAACAATCTC 1020
 Db 2299 TTAGACTTTCAGAGATTTATACGATCCCTCGTGATAAGCTAACTACTCTCAACAATCTC 2358
 Qy 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAGTAGAGGATAATACGATGAC 1080
 Db 2359 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAAAGTAGAGGATAATACGATGAC 2418
 Qy 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAT 1140
 Db 2419 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGAGATGCTAGCTAT 2478
 Qy 1141 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAAACGAGAAAGTTTACAGCTACTG 1200
 Db 2479 CATTTAGCTTATGATAAGATCGTTTATACCGAAGAGAAACGAGAAAGTTTACAGCTACTG 2538
 Qy 1201 CGTTATACAGGACACCTATACCTGATAACCCCTAAACGACAAATAA 1245
 Db 2539 CGTTATACAGGACACCTATACCTGATAACCCCTAAACGACAAATAA 2583
 RESULT 6
 AAQ12490
 ID AAQ12490 standard; DNA; 1467 BP.
 XX AC AAQ12490;
 XX AC
 XX 25-MAR-2003 (revised)
 DT 17-SEP-1991 (first entry)
 XX XX
 XX Factor Xa-cleavable streptokinase-IEGR-hirudin gene.
 XX Fusion protein; blood clotting; coagulation; fibrinolysis;
 KW antithrombotic; thrombolysis; ss.
 XX OS
 XX Synthetic.
 XX Key Location/Qualifiers
 FH 1. .1455
 PT CDS


```

FT mat_peptide /*tag= a
FT 1..1242
FT /*tag= b
FT /label= streptokinase
FT misc_RNA 1243..1254
FT /*tag= c
FT /label= linker
FT /note= "encodes factor Xa cleavage site"
FT mat_peptide 1255..1453
FT /*tag= d
FT /label= hirudin HV-1
FT
FT W09109125-A.
FT
FT 27-JUN-1991.
FT
FT 07-DEC-1989; 89GB-00027722.
FT
FT 07-DEC-1989; 89GB-00027722.
FT
FT 07-DEC-1990; 90WO-GB001911.
FT
FT (BRBT-) BRITISH BIO-TECHNOLOGY LTD.
FT
FT Dawson KM, Hunter MG, Czaplowski LG;
FT
FT WPI; 1991-208151/28.
FT
FT P-PSDB; AAR12522.
FT
FT Fusion protein cleavage by blood clotting enzyme - for prodn. of
FT fractions having greater antithrombotic activity for therapy and
FT prophylaxis.
FT
FT Disclosure; Page 98; 115pp; English.
FT
FT The sequence of the synthetic hirudin HV-1 gene was designed based on the
FT published amino acid sequence (Dodd J., et al FEBS letters 165:180
FT (1984)). The sequence of streptokinase was obtd. from PCR amplified
FT chromosomal DNA from S. equisimilis ATCC 10009 or ATCC 9642. The primers
FT used for the PCR were based on the pub- lished DNA sequence of S.
FT equisimilis strain H46A (Walke, H., Roe, B., and Ferretti, J.J., Gene 34
FT 357-362 (1985)). The two sequences were used to construct an expression
FT vector in which the streptokinase gene is linked to the hirudin gene via
FT a linking sequence encoding a cleavage site for factor Xa. The factor Xa
FT is present at the site of the target thrombus so the active agents are
FT released specifically at the place where clot formation is occurring. See
FT also AAQ12153-012156 and AAQ12158-012162. (Updated on 25-MAR-2003 to
FT correct PA field.)
FT
FT XX Sequence 1467 BP; 494 A; 317 C; 292 G; 364 T; 0 U; 0 Other;
FT
FT Query Match 99.5%; Score 1238.8; DB 2; Length 1467;
FT Best Local Similarity 99.8%; Pred. No. 0;
FT Matches 1240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
FT
FT 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATTAGTTGTT 60
FT |||||
FT 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCAAATTAGTTGTT 60
FT
FT 61 AGCGTTGCTGGTACTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGGAAATT 120
FT |||||
FT 61 AGCGTTGCTGGTACTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGGAAATT 120
FT
FT 121 GACCTAACATCAGCACTGCTCATGTAGGAGAAAGACAGACGAGCGTTAAGTCCAAATCA 180
FT |||||
FT 121 GACCTAACATCAGCACTGCTCATGTAGGAGAAAGACAGACGAGCGTTAAGTCCAAATCA 180
FT
FT 181 AAACCAATTTGCTACTGATAGTGGCGGATGCGGCATCAATCAATCTGAAAGCTGACTACTA 240
FT |||||
FT 181 AAACCAATTTGCTACTGATAGTGGCGGATGCGGCATCAATCAATCTGAAAGCTGACTACTA 240
FT
FT 241 AAGGCTATTCAAGAACAAATTTGATCGCTTAAAGTCCACAGTAACGCGACTACTTTGAGGTC 300
FT |||||
FT 241 AAGGCTATTCAAGAACAAATTTGATCGCTTAAAGTCCACAGTAACGCGACTACTTTGAGGTC 300

```

```

QY 301 ATTGATTTTGCAGGATGCAACCACTTACTGATCGAAACGGCAAGCTCTACTTTGCTGAC 360
|||
DB 301 ATTGATTTTGCAGGATGCAACCACTTACTGATCGAAACGGCAAGCTCTACTTTGCTGAC 360
QY 361 AAAGATGTTTCGGTAACTTGCAGCCCAACCTGTCCAAAGAAATTTTGTAAACGGACAT 420
|||
DB 361 AAAGATGTTTCGGTAACTTGCAGCCCAACCTGTCCAAAGAAATTTTGTAAACGGACAT 420
QY 421 GTGCGGTTAGACCATATAAAGAAAACCAATACAAATCAAGCGAAATCTGTGTGATG 480
|||
DB 421 GTGCGGTTAGACCATATAAAGAAAACCAATACAAATCAAGCGAAATCTGTGTGATG 480
QY 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTAGAGCAGGTCTCAAA 540
|||
DB 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTAGAGCAGGTCTCAAA 540
QY 541 GATATAAGCTATTGAAAACATAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600
|||
DB 541 GATATAAGCTATTGAAAACATAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600
QY 601 GCTCAGCACAAAGCATTTTTAAACCAACCCAGGCTATACGATTTATGAACTGAC 660
|||
DB 601 GCTCAGCACAAAGCATTTTTAAACCAACCCAGGCTATACGATTTATGAACTGAC 660
QY 661 TCCTCAATCGTCACTCATGCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 720
|||
DB 661 TCCTCAATCGTCACTCATGCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 720
QY 721 TTTATCTTACCATGTCAAAATCGGAAACAAAGCTTATGAGATCAATAAAAAATCTGCTG 780
|||
DB 721 TTTATCTTACCATGTCAAAATCGGAAACAAAGCTTATGAGATCAATAAAAAATCTGCTG 780
QY 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACCTCTCTTAAAAAGGG 840
|||
DB 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACCTCTCTTAAAAAGGG 840
QY 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 900
|||
DB 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 900
QY 901 GATGTCAACCAACCAAGTGTCTTAAAGGAGGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
|||
DB 901 GATGTCAACCAACCAAGTGTCTTAAAGGAGGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
QY 961 TTAGACTCTCAGAGATTTTATACGATCTCTGATGATGAGCTAAACCTCTCTACCAATCTC 1020
|||
DB 961 TTAGACTCTCAGAGATTTTATACGATCTCTGATGATGAGCTAAACCTCTCTACCAATCTC 1020
QY 1021 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAAAAGTAGAGGATATATCAAGATGAC 1080
|||
DB 1021 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAAAAGTAGAGGATATATCAAGATGAC 1080
QY 1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCGAAAGGAGAGATGCTAGCTAT 1140
|||
DB 1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCGAAAGGAGAGATGCTAGCTAT 1140
QY 1141 CATTAGCCCTATGATAAAGATCGTTTATACCGAAGAAAGACGAGAGTTTACAGCTACCTG 1200
|||
DB 1141 CATTAGCCCTATGATAAAGATCGTTTATACCGAAGAAAGACGAGAGTTTACAGCTACCTG 1200
QY 1201 CGTTATACAGGAGACCTATACCTGATTAACCTTAACGACAAA 1242
|||
DB 1201 CGTTATACAGGAGACCTATACCTGATTAACCTTAACGACAAA 1242

```

RESULT 7

AAA37633

ID AAA37633 standard; DNA; 1245 BP.

XX

AAA37633;

AC

XX 15-SEP-2003 (revised)

DT

DT 13-OCT-2000 (first entry)

XX S. equisimilis streptokinase coding sequence.

DE Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;

XX plasminogen; human; fibrinectin; thrombolytic therapy;

KW cardiovascular disorder; fibrinectin; ss.

XX Streptococcus dysgalactiae subsp. equisimilis.

OS

XX Key Location/Qualifiers

PH 1.1245

FT /*tag= a

FT /product= "streptokinase"

XX EPI024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99EP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

XX (COUL) CSIR COUNCIL SCI IND RES.

XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;

PI Yadav M;

XX WPI; 2000-516032/47.

DR P-PSDB; AAY90282.

XX Hybrid streptokinase-fibrin binding domain polypeptides useful for

PT thrombolytic therapy comprises a streptokinase fused with fibrin binding

PT domains of human fibrinectin.

XX Example 3; Fig 3; 58pp; English.

XX This sequence represents the human Streptococcus equisimilis

CC streptokinase coding sequence. The invention relates to a hybrid

CC plasminogen activator (PA) comprises a polypeptide fusion between

CC streptokinase (SK), which are capable of plasminogen (PG) activation, and

CC fibrin binding regions of human fibrinectin, which are from fibrin

CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the

CC ability to bind with fibrin independently and also characteristically

CC retains a PG activation ability which becomes evident only after a

CC pronounced duration, or lag, after exposure of the PA to a suitable

CC animal or human PG. The hybrid streptokinase-fibrin binding domain

CC polypeptides are useful in thrombolytic therapy for various kinds of

CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as

CC well as kinetics of plasminogen activation that are distinct from that of

CC natural streptokinase in being characterised by a temporary delay, or lag

CC of several minutes in the natural rate of the catalytic conversion of

CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins

CC can bind tightly with fibrin in blood clots soon after introduction into

CC the vascular system without significantly activating the circulating

CC blood plasminogen to plasmin, thus aiding in the localisation of the

CC plasminogen activation process to the site of pathological thrombus. This

CC overcomes systemic plasminogen activation encountered during clinical use

CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX

SQ Sequence 1245 BP; 426 A; 267 C; 237 G; 315 T; 0 U; 0 Other;

Query Match 98.3%; Score 1224.2; DB 3; Length 1245;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 1232; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 ATTGCTGACCTGATGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATTAGTTGTT 60

DB 1 ATTGCTGACCTGATGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATTAGTTGTT 60

QY 61 AGCGTTGCTGTTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTCGAATT 120

DB 61 AGCGTTGCTGTTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTCGAATC 120

121 GACCTAACATCAGCAGCTCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 180

121 GATCTAACATCAGCAGCTCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 180

181 AAACCAATTTGCTACTGATAGTGGGAGTCCACATAACTTGAAGAAAGCTGACTTACTA 240

181 AAACCAATTTGCTACTGATAGTGGGAGTCCACATAACTTGAAGAAAGCTGACTTACTA 240

241 AAGGCTATTCAAGCAAAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGTTC 300

241 AAGGCTATTCAAGCAAAATTTGATCGCTTAACGTCACAGTAACGACGACTACTTTGAGTTC 300

301 ATTGATTTTTCGAAGCGATGCAACCATTTACTGATCGAAACGCGAAGGCTCTACTTTGCTGAC 360

301 ATTGATTTTTCGAAGCGATGCAACCATTTACTGATCGAAACGCGAAGGCTCTACTTTGCTGAC 360

361 AAAGATGGTTGCGGTAACTTGGCGACCCCACTGTCCAGAAATTTTTCGTAAGCGGACAT 420

361 AAAGATGGTTGCGGTAACTTGGCGACCCCACTGTCCAGAAATTTTTCGTAAGCGGACAT 420

421 GTGCGGTTAGACCATATAAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480

421 GTGCGGTTAGACCATATAAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480

481 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 540

481 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 540

541 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 600

541 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 600

601 GCTCAAGCACAAGCATTTTTAAACAAACCCAGGCTATACGATTTATGAACGCTGAC 660

601 GCTCAAGCACAAGCATTTTTAAACAAACCCAGGCTATACGATTTATGAACGCTGAC 660

661 TCCTCAATGCTCTCATGACCAATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 720

661 TCCTCAATGCTCTCATGACCAATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 720

721 TTTACTTACCATGTCAAATTCGGGACCAAGCTTATGAGATCAATAAAATCTGGTCTG 780

721 TTTACTTACCATGTCAAATTCGGGACCAAGCTTATGAGATCAATAAAATCTGGTCTG 780

781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAGAGG 840

781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAGAGG 840

841 GAAAGCGGTATGATCCCTTTGATCGGACGTCATTGAAACTGTTCAAGATCAAAATAGTT 900

841 GAAAGCGGTATGATCCCTTTGATCGGACGTCATTGAAACTGTTCAAGATCAAAATAGTT 900

901 GATGTCACCAACCAAGATTTGCTAAAGCGAGCGCTCTTAAACGCTAGCGAACGTAAC 960

901 GATGTCACCAACCAAGATTTGCTAAAGCGAGCGCTCTTAAACGCTAGCGAACGTAAC 960

961 TTAGACTTCAGAGATTTTATACGATCCCTCGTGAAGGCTAAACTTCTACAAATCTC 1020

961 TTAGACTTCAGAGATTTTATACGATCCCTCGTGAAGGCTAAACTTCTACAAATCTC 1020

1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGGATATACGATGAC 1080

1021 GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAGGATATACGATGAC 1080

1081 ACCAACGTTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140

1081 ACCAACGTTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140

1141 CATTTAGCTATGATAAAGATCGTTTATACCGAAGAAAGAGAGGATTTTACAGCTACCTG 1200

1141 CATTTAGCTATGATAAAGATCGTTTATACCGAAGAAAGAGAGGATTTTACAGCTACCTG 1200

QY 1201 CGTTATACAGGACACCTATACCTGATACCCCTTAACGCAAAATAA 1245
DB 1201 CGTTATACAGGACACCTATACCTGATACCCCTTAACGCAAAATAA 1245

RESULT 8

AAQ05603
ID AAQ05603 standard; DNA; 1473 BP.

AC AAQ05603;

DT 25-MAR-2003 (revised)

DT 20-DEC-1990 (first entry)

XX Streptokinase G gene.

XX Streptokinase G; thrombolytic agent; plasminogen; plasmin;
KW recombinant plasmid; ss.

XX Streptococcus pyogenes.

XX Key Location/Qualifiers

FT CDS 60..1380

FT /*tag= a

FT /product= "streptokinase G"

XX DD276693-A.

XX 07-MAR-1990.

XX 07-NOV-1988; 88DD-00321531.

XX 07-NOV-1988; 88DD-00321531.

XX (DEAK) AKAD WISSENSCHAFTEN DDR.

XX Walter F, Siegel M, Malke H;

XX WPI; 1990-247327/33.

XX P-PSDB; AAR06377.

XX High yield streptokinase G prodn. from recombinant cells - transformed
PT with plasmid contg. gene from Streptococcus, useful as thrombolytic
PT agent.

XX Disclosure; Fig 2; 7pp; German.

XX Plasmid pMW1 (contg. the streptokinase G gene) is ligated with pUC19
CC and the resultant product used to transform bacteria. Infected cells are
CC cultured in liq. medium contg. assimilable C and N sources, and
CC streptokinase G is isolated from the cell lysate. The infected cells
CC provide the protein in high yields, i.e. 600 U/ml which is 1.5-2 times
CC that for the donor strain. See also DD-276694 (AAQ05604). (Updated on 25-
CC MAR-2003 to correct PA field.)

XX SQ Sequence 1473 BP; 498 A; 307 C; 272 G; 396 T; 0 U; 0 Other;

Query Match 98.3%; Score 1224.2; DB 2; Length 1473;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 1232; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGACCAATTAGTTGTT 60

DB 139 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGACCAATTAGTTGTT 198

QY 61 AGCGTTGCTGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTGAATTT 120

DB 199 AGCGTTGCTGTACTGTTGAGGGACGAATCAAGACATTAGTCTTAAATTTTGAATTT 258

QY 121 GACCTACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAGTCCCAATCA 180

DB 259 GACCTACATCAGACCTGCTCATGGAGGAAAGACAGACGAGCTTAAGTCCCAATCA 318

QY 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAAGAGCTGCTACTA 240

DB 319 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAAGAGCTGCTACTA 378

QY 241 AAGGCTATTTCAAGAAACAATTGATTCGCTAACCGTCCACAGTAAACGACACTACTTTGAGGTC 300

DB 379 AAGGCTATTTCAAGAAACAATTGATTCGCTAACCGTCCACAGTAAACGACACTACTTTGAGGTC 438

QY 301 ATTGATTTTTCGAAGCGATGCAACCACTTACTGATCGAAACGGCAAGGTCTTACTTTGCTGAC 360

DB 439 ATTGATTTTTCGAAGCGATGCAACCACTTACTGATCGAAACGGCAAGGTCTTACTTTGCTGAC 498

QY 361 AAAGATGGTTTCGGTAACTTCGCGACCCCAACCTGTCCAAAGAAATTTTGTAAAGCGGACAT 420

DB 499 AAAGATGGTTTCGGTAACTTCGCGACCCCAACCTGTCCAAAGAAATTTTGTAAAGCGGACAT 558

QY 421 GTGCGGTTAGACCATATATAAGAAACCAATACAAAATCAAGCGAAATCTGTGTATGTG 480

DB 559 GTGCGGTTAGACCATATATAAGAAACCAATACAAAATCAAGCGAAATCTGTGTATGTG 618

QY 481 GAATATACGTACAGTTTACTCCCTTAAACCTGTGATGACGATTTTCAAGCCAGGTCTCAAA 540

DB 619 GAATATACGTACAGTTTACTCCCTTAAACCTGTGATGACGATTTTCAAGCCAGGTCTCAAA 678

QY 541 GATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAATA 600

DB 679 GATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAATA 738

QY 601 GCTCAAGCAACAAGCAATTTTAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 660

DB 739 GCTCAAGCAACAAGCAATTTTAAACCAACCCAGGCTATACGATTTTATGAACGTGAC 798

QY 661 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720

DB 799 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 858

QY 721 TTTACTTACCATGTCAAAAATCGGAAACAAAGCTTATGAGATCAATAAAAAATCTGTCGTCG 780

DB 859 TTTACTTACCATGTCAAAAATCGGAAACAAAGCTTATGAGATCAATAAAAAATCTGTCGTCG 918

QY 781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840

DB 919 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 978

QY 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAATACGTT 900

DB 979 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCCACCATCAATACGTT 1038

QY 901 GATGTCAACACCAACGAATTTGCTAAAAAGGAGAGAGCTCTTAAACAGCTAGCGAACGTAAC 960

DB 1039 GATGTCAACACCAACGAATTTGCTAAAAAGGAGAGAGCTCTTAAACAGCTAGCGAACGTAAC 1098

QY 961 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGCTAAACTACTCTCAACAATCTC 1020

DB 1099 TTAGACTTCAGAGATTTATACGATCCCTCGTGATAAGCTAAACTACTCTCAACAATCTC 1158

QY 1021 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAAAAGTAGAGGATATCAACGATGAC 1080

DB 1159 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAAAAGTAGAGGATATCAACGATGAC 1218

QY 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140

DB 1219 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1278

QY 1141 CATTTAGCTTATGATAAAGATCGTTATACCGAAGAGACGAGAAGTTTACAGCTACCTG 1200

DB 1279 CATTTAGCTTATGATAAAGATCGTTATACCGAAGAGACGAGAAGTTTACAGCTACCTG 1338

QY 1201 CGTTATACAGGACACCTTATACCTGATAACCCCTTAACGCAAAATAA 1245

DB 1339 CGTTATACAGGACACCTTATACCTGATAACCCCTTAACGCAAAATAA 1383

RESULT 9
AAA37622
ID AAA37622 standard; DNA; 1377 BP.
XX AC
XX AAA37622;
XX AC
DT 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX ST
XX Streptokinase-NTRN gene.
XX ST
XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX ST
XX Streptococcus dysgalactiae subsp. equisimilis.
OS
XX
XX EP1024192-A2.
XX
XX 02-AUG-2000.
XX PF
XX 23-DEC-1999; 99EP-00310541.
XX PR
XX 24-DEC-1998; 98IN-DE003825.
XX
XX (COUL) CSIR COUNCIL SCI IND RES.
XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX
XX WPI; 2000-516032/47.
XX
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.
XX
XX Example 1; Fig 11; 58pp; English.
XX
XX This sequence represents a streptokinase-NTRN (SK-NTRN) gene (where NTRN
CC stands for N-terminally repaired with native sequence). The invention
CC relates to a hybrid plasminogen activator (PA) comprises a polypeptide
CC fusion between streptokinase (SK), which are capable of plasminogen (PG)
CC activation, and fibrin binding regions of human fibrinectin, which are
CC from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA
CC possesses the ability to bind with fibrin independently and also
CC characteristically retains a PG activation ability which becomes evident
CC only after a pronounced duration, or lag, after exposure of the PA to a
CC suitable animal or human PG. The hybrid streptokinase-fibrin binding
CC domain polypeptides are useful in thrombolytic therapy for various kinds
CC of cardiovascular disorders. The hybrids have enhanced fibrin selectivity
CC as well as kinetics of plasminogen activation that are distinct from that
CC of natural streptokinase in being characterised by a temporary delay, or
CC lag of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 1377 BP; 465 A; 302 C; 265 G; 345 T; 0 U; 0 Other;
Query Match 98.2%; Score 1222.6; DB 3; Length 1377;
Best Local Similarity 98.9%; Pred. NO. 0;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACACAGCCCAATGTTGTT 60
Db 133 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCACACAGCCCAATGTTGTT 192
Qy 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATTTTTTGAATT 120

Db 193 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAATTTTTTGAATC 252
Qy 121 GACCTAACATCACGACCTGCTCATGAGGAGAAAGACAGAGCAAGCGCTTAAGTCCAAATCA 180
Db 253 GATCTAACATCACGACCTGCTCATGAGGAGAAAGACAGAGCAAGCGCTTAAGTCCAAATCA 312
Qy 181 AAACCAATTTGCTACTGATGATGCGCGGATGCCACATAAACTTGAAAAGCTGACTTACTA 240
Db 313 AAACCAATTTGCTACTGATGATGCGCGGATGCCACATAAACTTGAAAAGCTGACTTACTA 372
Qy 241 AAGGCTATTCAAGAACAAATTTGATGCTTAAGCTCCACAGTAACGACGACTTCTTTGAGGTC 300
Db 373 AAGGCTATTCAAGAACAAATTTGATGCTTAAGCTCCACAGTAACGACGACTTCTTTGAGGTC 432
Qy 301 ATTGATTTTGCAGAGCGATGCAACATTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 360
Db 433 ATTGATTTTGCAGAGCGATGCAACATTACTGATCGAAACGGCAAGGCTCTACTTTGCTGAC 492
Qy 361 AAGATGTTTTCGGTAACTTCCGACCCCACTGTCGAGAAATTTTTCGCTAAGCGGCAT 420
Db 493 AAGATGTTTTCGGTAACTTCCGACCCCACTGTCGAGAAATTTTTCGCTAAGCGGCAT 552
Qy 421 GTGCGGTTAGACCATATAAAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
Db 553 GTGCGGTTAGACCATATAAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 612
Qy 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTTCTCAA 540
Db 613 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGTTCTCAA 672
Qy 541 GATACTAAGCTATTGAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 600
Db 673 GATACTAAGCTATTGAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 732
Qy 601 GCTCAAGCACAAAGCATTTTAAACAAAACCCAGGCTATACGATTTATGAACGTTGAC 660
Db 733 GCTCAAGCACAAAGCATTTTAAACAAAACCCAGGCTATACGATTTATGAACGTTGAC 792
Qy 661 TCCTCAATGCTACTCANGACAATGACATTTTCGGTAGATTTTACCAATGGATCAAGAG 720
Db 793 TCCTCAATGCTACTCANGACAATGACATTTTCGGTAGATTTTACCAATGGATCAAGAG 852
Qy 721 TTTTACTTACCATCTCAAAATCGGACACAGCTTATGAGATCAATAAAAATCTGGTCTG 780
Db 853 TTTTACTTACCATCTCAAAATCGGACACAGCTTATGAGATCAATAAAAATCTGGTCTG 912
Qy 781 AATGAAGAAATAAACCAACACTGACCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
Db 913 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 972
Qy 841 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACCTGAACTGTTCCACATCAATAGTT 900
Db 973 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACCTGAACTGTTCCACATCAATAGTT 1032
Qy 901 GATGTCAACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTACGACACGTAAC 960
Db 1033 GATGTGATACCAACGAATTTGCTAAAAAGTGACAGCTCTTAAACAGCTACGACACGTAAC 1092
Qy 961 TTAGACTTCAGAGATTTTATACGATCCCTCGTGAAGCTAAACTTCTCTCAACAATCTC 1020
Db 1093 TTAGACTTCAGAGATTTTATACGATCCCTCGTGAAGCTAAACTTCTCTCAACAATCTC 1152
Qy 1021 GATGCTTTTGGTATTATGGAATTAACCTTAACCTGGAAGAGTAGAGGATTAATCAGATGAC 1080
Db 1153 GATGCTTTTGGTATTATGGAATTAACCTTAACCTGGAAGAGTAGAGGATTAATCAGATGAC 1212
Qy 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCAGCCCGAAGGAGAGAGATGCTAGCTAT 1140
Db 1213 ACCAACCGTATCATACCGTTTATATGGGCAAGCAGCCCGAAGGAGAGAGATGCTAGCTAT 1272
Qy 1141 CATTTAGSCCTATGATAAAGATCGTTTATACCGAAGAGAAAGAGAGATTTTACGACTCCTG 1200
Db 1273 CATTTAGSCCTATGATAAAGATCGTTTATACCGAAGAGAAAGAGAGATTTTACGACTCCTG 1332

QY 1201 CGTTATACAGGACACCTATACCTGATACCTTAACCCCTAACGACAAATAA 1245
Db 1333 CGTTATACAGGACACCTATACCTGATACCTTAACCCCTAACGACAAATAA 1377

RESULT 10
AAA37642
ID AAA37642 standard; DNA; 1782 BP.
XX AAA37642;
AC
XX
DT 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX Chimeric SK-FBD coding sequence.
XX
KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibronectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.
XX
PN EP1024192-A2.
XX
PD 02-AUG-2000.
XX
PF 23-DEC-1999; 99BP-00310541.
XX
PR 24-DEC-1998; 98IN-DR003825.
XX (COUL) CSIR COUNCIL SCI IND RES.
XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX
XX WPI; 2000-516032/47.
XX
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibronectin.
XX
XX Example 5; Fig 21b; 58pp; English.
XX
CC This sequence represents a chimeric streptokinase-fibrin binding domain
CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibronectin, which are from fibrin
CC binding domains (PBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 1782 BP; 573 A; 407 C; 377 G; 425 T; 0 U; 0 Other;

Query Match 98.2%; Score 1222.6; DB 3; Length 1782;
Best Local Similarity 98.9%; Pred. No. 0;

Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTAGTTGT 60
Db 538 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAACAGCAATTTAGTTGT 597

QY 61 AGCGTTGCTGTGTAATCTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAAT 120
Db 598 AGCGTTGCTGTGTAATCTGTTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTGAAT 657

QY 121 GACCTAAATCAAGACCTGCTCATGGAGGAAGACAGACAGAGGCTTAATGCCAAATCA 180
Db 658 GATCTAAACATCAAGACCTGCTCATGGAGGAAGACAGACAGAGGCTTAATGCCAAATCA 717

QY 181 AAACCAATTTGCTACTGATAGTGGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA 240
Db 718 AAACCAATTTGCTACTGATAGTGGCGATGCCACATAAACTTGAAAAAGCTGACTTACTA 777

QY 241 AAGGCTATTCAAGAACAAATTTGCTAAAGCTCCACAGTAACGACGACTTCTTGAAGTTC 300
Db 778 AAGGCTATTCAAGAACAAATTTGCTAAAGCTCCACAGTAACGACGACTTCTTGAAGTTC 837

QY 301 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAGCGCAAGCTCTACTTTGCTGAC 360
Db 838 ATTGATTTTGAAGCGATGCAACCATTTACTGATCGAAGCGCAAGCTCTACTTTGCTGAC 897

QY 361 AAAGATGGTTTCGTTAACTTTCGCGACCCCAACCTGTCCAAAGAAATTTTGTGTAAGCGGACAT 420
Db 898 AAAGATGGTTTCGTTAACTTTCGCGACCCCAACCTGTCCAAAGAAATTTTGTGTAAGCGGACAT 957

QY 421 GTGCGGTTAGACCATATATAAGAAAAACCAATATAAAATCAAGCGAATCTGTTGATGTG 480
Db 958 GTGCGGTTAGACCATATATAAGAAAAACCAATATAAAATCAAGCGAATCTGTTGATGTG 1017

QY 481 GAATATCTGTACAGTTTACTCCCTTAAACCTGTGATGAGGATTTTCAGACAGGCTCTCAA 540
Db 1018 GAATATCTGTACAGTTTACTCCCTTAAACCTGTGATGAGGATTTTCAGACAGGCTCTCAA 1077

QY 541 GATCTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCACTCTCAAGAAATTTACTA 600
Db 1078 GATCTAAGCTATTGAAAAACATAGCTATCGGTGACACCATCACTCTCAAGAAATTTACTA 1137

QY 601 GGTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGAACGTCAC 660
Db 1138 GGTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTATACGATTTATGAACGTCAC 1197

QY 661 TCCTCAATCGTCACTCATGACATGACATTTTCCTGACGATTTTACCAATGGATCAAGAG 720
Db 1198 TCCTCAATCGTCACTCATGACATGACATTTTCCTGACGATTTTACCAATGGATCAAGAG 1257

QY 721 TTTACTTACCATGTCAAAAATCGGAAACAAAGCTTATGAGATCAATAAAAAATCTGCTG 780
Db 1258 TTTACTTACCATGTCAAAAATCGGAAACAAAGCTTATGAGATCAATAAAAAATCTGCTG 1317

QY 781 AATGAGAAATTAACAACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 840
Db 1318 AATGAGAAATTAACAACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1377

QY 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCCACCATCAATAGCTT 900
Db 1378 GAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCCACCATCAATAGCTT 1437

QY 901 GATGTCAACCAACAAATGCTTAAAGAGGAGCAGCTCTTAAACAGTGGAGCAAGTAAC 960
Db 1438 GATGTGTATCAACCAACAAATGCTTAAAGAGGAGCAGCTCTTAAACAGTGGAGCAAGTAAC 1497

QY 961 TTAGACTTCAGAGATTTTACGATCCCTGATGTAAGCTTAACTCTCTACACAATCTC 1020
Db 1498 TTAGACTTCAGAGATTTTACGATCCCTGATGTAAGCTTAACTCTCTACACAATCTC 1557

QY 1021 GATGCTTTTGGTATTATGACTATATCTTAACTGGAAGAGTAGAGGATATCAAGATGAC 1080
Db 1558 GATGCTTTTGGTATTATGACTATATCTTAACTGGAAGAGTAGAGGATATCAAGATGAC 1617

1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGAATGCTAGCTAT 1140
 1618 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGAATGCTAGCTAT 1677
 1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1200
 1678 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTACCTG 1737
 1201 CGTTATACAGGACACCTATACCTGATACCTTAACGACGAATAA 1245
 1738 CGTTATACAGGACACCTATACCTGATACCTTAACGACGAATAA 1782

RESULT 11
 AAQ11651
 ID AAQ11651 standard; DNA; 2030 BP.
 XX
 AC AAQ11651;
 XX
 DT 08-JUL-1991 (first entry)
 XX
 DE FB-FB-SK fusion construct.
 XX
 KW Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
 KW streptokinase; fusion protein; ss.
 XX
 OS Staphylococcus aureus.
 XX
 FH Key Location/Qualifiers
 FT CDS 10..184
 FT /*tag= a
 FT /label= FB monomer
 FT CDS 185..358
 FT /*tag= b
 FT /label= FB monomer
 FT CDS 359..1601
 FT /*tag= c
 FT /label= streptokinase
 XX
 PN US5011686-A.
 XX
 PD 30-APR-1991.
 XX
 PF 15-NOV-1989; 89US-00437769.
 XX
 PR 21-SEP-1987; 87US-00099242.
 XX
 PA (CREA-) CREATIVE BIOMOLEC.
 XX
 PI Pang RHL;
 XX
 DR WPI; 1991-140198/19.
 DR P-PSDB; AAR11829.
 XX
 PT Imparting injectable fibrinolytic agent - with affinity for intravascular
 PT thrombus, by linking agent to fibrin binding domain.
 XX
 PS Disclosure; Fig 5; 18pp; English.
 XX
 CC The DNA encodes an FB-FB dimer linked to the streptokinase coding
 CC sequence. The FB fragment has selective affinity for fibrin, low affinity
 CC for fibrinogen, and minimal immunogenicity, imparting thrombus-targeting
 CC capability. See also AAQ11649 and AAQ11650
 XX
 SQ Sequence 2030 BP; 673 A; 464 C; 406 G; 487 T; 0 U; 0 Other;
 Query Match 98.2%; Score 1222.6; DB 2; Length 2030;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 1 ATTGCTGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACACGCAATAGTTGTT 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 358 ATTGCTGACCTGAGTGGCTCTAGACCGTCCATCTGTCAACACACCAATAGTTGTT 417
 QY 61 AGGTTTGTCTGCTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTTTGAAT 120
 Db 418 AGGTTTGTCTGCTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTTTGAAT 477
 QY 121 GACCTAACATCAAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 180
 Db 478 GATCTAACATCAAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 537
 QY 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAATTTGAAAAGCTGACTACTA 240
 Db 538 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAATTTGAAAAGCTGACTACTA 597
 QY 241 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAAACGACACTACTTTGAGGTC 300
 Db 598 AAGGCTATTCAAGAACAAATTTGATCGCTTAACGCTCCACAGTAAACGACACTACTTTGAGGTC 657
 QY 301 ATTGATTTTCAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 360
 Db 658 ATTGATTTTCAAGCGATGCAACCATTTACTGATCGAAACGGCAAGGCTACTTTGCTGAC 717
 QY 361 AAGATGGTTCGTAACCTTGGCGACCCACCTGTCGAGAAATTTTGTAAAGGACAT 420
 Db 718 AAGATGGTTCGTAACCTTGGCGACCCACCTGTCGAGAAATTTTGTAAAGGACAT 777
 QY 421 GTGCGGTTAGACCATATAAAGAAAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480
 Db 778 GTGCGGTTAGACCATATAAAGAAAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 837
 QY 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 540
 Db 838 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 897
 QY 541 GATACTAAGCTATTGAAACACATAGTATCGGTGACACCATCATCTCAAGAAATTAATA 600
 Db 898 GATACTAAGCTATTGAAACACATAGTATCGGTGACACCATCATCTCAAGAAATTAATA 957
 QY 601 GCTCAAGCACAAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACGCTGAC 660
 Db 958 GCTCAAGCACAAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACGCTGAC 1017
 QY 661 TCCTCAATCGTACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 720
 Db 1018 TCCTCAATCGTACTCATGACATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 1077
 QY 721 TTTTACTTACCATGTCAAAAATCGGAACAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
 Db 1078 TTTTACTTACCATGTCAAAAATCGGAACAGCTTATGAGATCAATAAAAAATCTGGTCTG 1137
 QY 781 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGCTTAAAAAGGG 840
 Db 1138 AATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTACGCTTAAAAAGGG 1197
 QY 841 GAAAGCGGTATGATCCCTTTGATCCGAGTCACTTGAACCTGTTTACCAATCAATAGCTT 900
 Db 1198 GAAAGCGGTATGATCCCTTTGATCCGAGTCACTTGAACCTGTTTACCAATCAATAGCTT 1257
 QY 901 GATGTCAACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGTACGCAACGTAAC 960
 Db 1258 GATGTCAACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGTACGCAACGTAAC 1317
 QY 961 TTAGACTTCAGAGATTTTATACGATCCCTGATGTAAGGCTAAACTCTCTCAACAAATCTC 1020
 Db 1318 TTAGACTTCAGAGATTTTATACGATCCCTGATGTAAGGCTAAACTCTCTCAACAAATCTC 1377
 QY 1021 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGATAGGAGATATACGATGAC 1080
 Db 1378 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGATAGGAGATATACGATGAC 1437
 QY 1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGAATGCTAGCTAT 1140
 Db 1438 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGAGAGAATGCTAGCTAT 1497

QY 1141 CATTAGCCTATGATAAAGATCGTTATACCGAAGAAACGAGAGCTTTTACAGCTACCTG 1200
Db 1498 CATTAGCCTATGATAAAGATCGTTATACCGAAGAAACGAGAGCTTTTACAGCTACCTG 1557
QY 1201 CATTAGCCTATGATAAAGATCGTTATACCGAAGAAACGAGAGCTTTTACAGCTACCTG 1245
Db 1558 CATTAGCCTATGATAAAGATCGTTATACCGAAGAAACGAGAGCTTTTACAGCTACCTG 1602

RESULT 12
ID AAX80492
XX AAX80492 standard; cDNA; 1242 BP.
AC AAX80492;
XX 17-OCT-2003 (revised)
DT 26-AUG-1999 (first entry)
XX Streptococcus equisimilis native streptokinase encoding cDNA.
XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW nSK; rSK; bacterial; blood clot; thrombotic condition;
KW myocardial infarction; venous thrombosis; pulmonary embolism;
KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
XX Streptococcus dysgalactiae subsp. equisimilis.
OS
XX WO9931247-A1.
XX 24-JUN-1999.
XX 15-DEC-1998; 98WO-US026694.
XX 15-DEC-1997; 97US-0069497P.
PR (HARD) HARVARD COLLEGE.
XX Reed GL;
XX WPI; 1999-395183/33.
DR P-FSDB; AAY24794.
XX N-terminally deleted streptokinase.
XX Claim 44; Page 58-60; 73pp; English.
XX The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase, the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase protein; a
CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (1); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising a
CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
CC blood clots in patients with a thrombotic condition, e.g. myocardial
CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
CC graft thrombosis and arterial thrombosis. The modified streptokinase can
CC also be used in non-human mammals. Streptokinase activation of
CC plasminogen is at least 10-fold, preferably 100-fold greater in the
CC presence of fibrin than in the absence of fibrin. The modified
CC streptokinase has at least one amino acid substitution that inactivates a
CC substrate site for proteolytic cleavage. This reduces the rate of
CC degradation of the streptokinase at least two-fold. The present sequence
CC encodes native streptokinase (nSK). (Updated on 17-OCT-2003 to
XX standardise OS field)
SQ Sequence 1242 BP; 424 A; 267 C; 237 G; 314 T; 0 U; 0 Other;
Query Match 98.1%; Score 1221.2; DB 2; Length 1242;

Best Local Similarity 99.0%; Pred. No. 0;
Matches 1229; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCAATTTAGTTGTT 60
Db 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAACGCAATTTAGTTGTT 60
QY 61 AGCGTTGCTGTGTTGAGGGGACGAATCAAGACATTAGTCTTTAAATTTTTTGAAT 120
Db 61 AGCGTTGCTGTGTTGAGGGGACGAATCAAGACATTAGTCTTTAAATTTTTTGAAT 120
QY 121 GACCTAAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 180
Db 121 GATCTAAACATCAAGACCTGCTCATGAGGAAAGACAGCAAGGCTTAAGTCCAAATCA 180
QY 181 AAACCATTTGCTACTGATAGTGGCGGATGCCACATATAAACTTGAAAAAGCTGACTTACTA 240
Db 181 AAACCATTTGCTACTGATAGTGGCGGATGCCACATATAAACTTGAAAAAGCTGACTTACTA 240
QY 241 AAGGCTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAAACGACACTCTTTGAGGTC 300
Db 241 AAGGCTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAAACGACACTCTTTGAGGTC 300
QY 301 ATTGATTTGCAAGCATGCAACCATTTACTGATCGAAACGGCAAGTCTTACTTTGCTGAC 360
Db 301 ATTGATTTGCAAGCATGCAACCATTTACTGATCGAAACGGCAAGTCTTACTTTGCTGAC 360
QY 361 AAAGATGTTTCGTTAACCTTGGCGACCCCACTGTCCAAGAAATTTTGTCTAAGCGGACAT 420
Db 361 AAAGATGTTTCGTTAACCTTGGCGACCCCACTGTCCAAGAAATTTTGTCTAAGCGGACAT 420
QY 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
Db 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
QY 481 GAATATCTGTACAGTTTACTCCCTTAAACCTGTGATGAGATTTTCAGACAGCTCTCAAA 540
Db 481 GAATATCTGTACAGTTTACTCCCTTAAACCTGTGATGAGATTTTCAGACAGCTCTCAAA 540
QY 541 GATCTAAGCTATTGAAAAACACATAGCTATCGGTGACACCATCACATCTCAAGAAATTA 600
Db 541 GATCTAAGCTATTGAAAAACACATAGCTATCGGTGACACCATCACATCTCAAGAAATTA 600
QY 601 GCTCAAGCAAAAGCAATTTTAAACCAAAACCCAGGCTATACGATTTTATGAACGTCAC 660
Db 601 GCTCAAGCAAAAGCAATTTTAAACCAAAACCCAGGCTATACGATTTTATGAACGTCAC 660
QY 661 TCCTCAATCGTCACTCATGACATGACATTTTCCTGACGATTTTACCAATGGATCAAGAG 720
Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCCTGACGATTTTACCAATGGATCAAGAG 720
QY 721 TTTTACTTACCATGTCAAAATCGGAAACAGCTTTATGAGATCAATAAAAAATCTGGTCTG 780
Db 721 TTTTACTTACCATGTCAAAATCGGAAACAGCTTTATGAGATCAATAAAAAATCTGGTCTG 780
QY 781 ATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATTTATGCTCTTAAAAAGGG 840
Db 781 ATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATTTATGCTCTTAAAAAGGG 840
QY 841 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTTCAACCTCAATACGTT 900
Db 841 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTTCAACCTCAATACGTT 900
QY 901 GATGTCAACACCAACCAATTTGCTAAAAAGCGAGCTCTTAAACAGTAGCGCAACGTAAC 960
Db 901 GATGTCAACACCAACCAATTTGCTAAAAAGCGAGCTCTTAAACAGTAGCGCAACGTAAC 960
QY 961 TTAGCTTCAGAGATTTTACGATCTCTGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 961 TTAGCTTCAGAGATTTTACGATCTCTGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 GATGCTTTTGGTATTATGAGTATATCTTAACTGGAAGAGTAGAGGATTAATCAGGATGAC 1080

Db 1021 GATGCTTTTGGTATTATGAGTACTATACCTTAACCTGAAAGAGTAGAGGATAATCAGTAGAC 1080
 QY 1081 ACCAACGGTATCATACCGTTTATATGCGCAAGCGACCCGAGGAGAGAGATGCTAGCTAT 1140
 Db 1081 ACCAACGGTATCATACCGTTTATATGCGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
 QY 1141 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAGAGAGAGATGCTAGCTAT 1200
 Db 1141 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAGAGAGAGAGATGCTAGCTAT 1200
 QY 1201 CGTTATACAGGACACCTATACCTGATACCTTAACCTTAACGACAAA 1242
 Db 1201 CGTTATACAGGACACCTATACCTGATACCTTAACCTTAACGACAAA 1242

RESULT 13

ABAU5546
 ID ABA05546 standard; cDNA; 1254 BP.

XX ABA05546;
 DT 26-FEB-2002 (first entry)
 XX Streptokinase cDNA.

XX Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;
 KW vasodilator; thrombolytic; angina; myocardial infarction; stroke;
 KW gene therapy; maxadilan; ss.
 XX Unidentified.

XX WO200185100-A2.

XX 15-NOV-2001.

XX 10-MAY-2001; 2001WO-US015209.

XX 11-MAY-2000; 2000US-00569920.

XX (GEO) GEN HOSPITAL CORP.

XX Reddy VB, Lerner E;

XX WPI; 2002-062184/08.

XX New fusion protein or conjugate, useful for treating unstable angina,
 PT acute myocardial infarction or stroke, comprises a vasodilator
 PT polypeptide and a thrombolytic polypeptide, or active fragments of the
 PT polypeptides.

XX Example 1; Fig 2; 37pp; English.

XX The invention relates to a fusion protein or a conjugate comprising a
 CC vasodilator polypeptide, or its active fragment, and a thrombolytic
 CC polypeptide or its active fragment. The protein is useful for treating a
 CC subject suffering from a partially or totally occluded blood vessel,
 CC causing unstable angina, acute myocardial infarction or stroke. The
 CC polynucleotide encoding the polypeptide is useful in gene therapy. The
 CC vasodilatory action of the protein allows for the use of lower doses of a
 CC thrombolytic while maintaining the clot dissolving effectiveness of the
 CC thrombolytic, and the use of lower doses of the thrombolytic reduces
 CC associated side effects. The present sequence is the streptokinase cDNA
 CC used in the construction of a Maxadilan-Streptokinase fusion protein.
 CC Maxadilan is a vasodilator peptide produced by the salivary gland of the
 CC New World sand fly

XX Sequence 1254 BP; 426 A; 271 C; 241 G; 316 T; 0 U; 0 Other;

XX Query Match 98.1%; Score 1221.2; DB 6; Length 1254;

XX Best Local Similarity 99.0%; Pred. No. 0;

XX Matches 1229; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCACACAGGCAATTAGTTGTT 60

Db 7 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCACACAGGCAATTAGTTGTT 66
 QY 61 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAAT 120
 Db 67 AGCGTTGCTGGTACTGTTGAGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAAT 126
 QY 121 GACCTAACATCAACGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 180
 Db 127 GATCTAACATCAACGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 186
 QY 181 AAACCATTTGCTACTGATAGTGGCGGATCCACATTAATCTTGAAAAGCTGACTACTA 240
 Db 187 AAACCATTTGCTACTGATAGTGGCGGATCCACATTAATCTTGAAAAGCTGACTACTA 246
 QY 241 AAGGCTATTCAAGAAACAATGATCGCTTAACGCTCCACAGTAAACGACGACTACTTTGAGGTC 300
 Db 247 AAGGCTATTCAAGAAACAATGATCGCTTAACGCTCCACAGTAAACGACGACTACTTTGAGGTC 306
 QY 301 ATTGATTTTCAAAGCGATGCAACATTAATGATGCGAAACCGGCAAGGCTACTTTGCTGAC 360
 Db 307 ATTGATTTTCAAAGCGATGCAACATTAATGATGCGAAACCGGCAAGGCTACTTTGCTGAC 366
 QY 361 AAAGATGGTTCGGTAACTTGGCGGACCCCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 420
 Db 367 AAAGATGGTTCGGTAACTTGGCGGACCCCAACCTGTCCAGAAATTTTGTCTAAGCGGACAT 426
 QY 421 GTGCGGCTTAGAGCCATATAAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480
 Db 427 GTGCGGCTTAGAGCCATATAAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 486
 QY 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 540
 Db 487 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 546
 QY 541 GATTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 600
 Db 547 GATTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTA 606
 QY 601 GCTCAAGCAACAAAGCATTTTAAACAAACCCACCCAGGCTATACGATTTATGAACGTCAC 660
 Db 607 GCTCAAGCAACAAAGCATTTTAAACAAACCCACCCAGGCTATACGATTTATGAACGTCAC 666
 QY 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTAGGATTTTACCAATGGATCAAGAG 720
 Db 667 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTAGGATTTTACCAATGGATCAAGAG 726
 QY 721 TTTACTTACCATGTCAAAATCGSGAAACAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
 Db 727 TTTACTTACCGTGTAAAAATCGSGAAACAGCTTATGAGATCAATAAAAAATCTGGTCTG 786
 QY 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 840
 Db 787 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 846
 QY 841 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACATCAATAGCTT 900
 Db 847 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACATCAATAGCTT 906
 QY 901 GATGTCACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 960
 Db 907 GATGTCACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAACGTAAC 966
 QY 961 TTAGACTTCAGAGATTTTATACGATCTCTGTTGATAAGGCTAAATCTCTCTACCAATCTC 1020
 Db 967 TTAGACTTCAGAGATTTTATACGATCTCTGTTGATAAGGCTAAATCTCTCTACCAATCTC 1026
 QY 1021 GATGCTTTTGGTATTATGAGTACTATACCTTAACCTGAAAAGTAGAGGATTAATCAGGATGAC 1080
 Db 1027 GATGCTTTTGGTATTATGAGTACTATACCTTAACCTGAAAAGTAGAGGATTAATCAGGATGAC 1086
 QY 1081 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCCGAGGAGAGAGATGCTAGCTAT 1140

Db 1087 ACCAACCGTATCATACACCGTTTATATGGGCAAGCGACCCGGAAGAGAGAAATGCTAGTAT 1146
QY 1141 CATTTAGCCCTATGATAAAGATCGTTTACCGAAGAAGAACGAGAGCTTTTACAGCTACCTG 1200
Db 1147 CATTTAGCCCTATGATAAAGATCGTTTACCGAAGAAGAACGAGAGCTTTTACAGCTACCTG 1206
QY 1201 CGTTATACAGGGACACCTATACCTGATAACCCCTAACGACAAA 1242
Db 1207 CGTTATACAGGGACACCTATACCTGATAACCCCTAACGACAAA 1248

RESULT 14

AXX80497
ID AAX80497 standard; cDNA; 2385 BP.
XX
AC AAX80497;
XX
XX 26-AUG-1999 (first entry)
XX Streptokinase and maltose binding protein fusion protein encoding cDNA.
DE Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW rSK; rSK; Bacterial; blood clot; thrombotic condition;
KW myocardial infarction; venous thrombosis; pulmonary embolism;
KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.
XX Streptococcus dysgalactiae subsp. equisimilis.
OS Synthetic.
OS
XX W09931247-A1.
FN
XX 24-JUN-1999.
PD
XX 15-DEC-1998; 98WO-US026694.
XX
XX 15-DEC-1997; 97US-0069497P.
PR (HARD) HARVARD COLLEGE.
XX Reed GL;
XX
XX WPI; 1999-395183/33.
DR P-PSDB; AAY24797.
XX
XX N-terminally deleted streptokinase.
XX
XX Example; Page 45-48; 73pp; English.
XX
CC The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase, the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase protein; a
CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (1); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising a
CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
CC blood clots in patients with a thrombotic condition, e.g. myocardial
CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
CC graft thrombosis and arterial thrombosis. The modified streptokinase can
CC also be used in non-human mammals. Streptokinase activation of
CC plasminogen is at least 10-fold, preferably 100-fold greater in the
CC presence of fibrin than in the absence of fibrin. The modified
CC streptokinase has at least one amino acid substitution that inactivates a
CC substrate site for proteolytic cleavage. This reduces the rate of
CC degradation of the streptokinase at least two-fold. The present sequence
CC encodes a streptokinase and maltose binding protein fusion protein from
CC an example of the present invention
XX
XX Sequence 2385 BP; 745 A; 547 C; 549 G; 544 T; 0 U; 0 Other;

Query Match 98.1%; Score 1221.2; DB 2; Length 2385;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1229; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAGACCCAAATTAGTTGTT 60
Db 1144 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAGACCCAAATTAGTTGTT 1203
QY 61 AGCGTTGCTGTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 120
Db 1204 AGCGTTGCTGTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAAT 1263
QY 121 GACCTAACATCAACGACCTGCTCATGAGGAGAAAGACAGACAGGCTTAAAGTCCAAATCA 180
Db 1264 GATCTAAACATCAACGACCTGCTCATGAGGAGAAAGACAGACAGGCTTAAAGTCCAAATCA 1323
QY 181 AAACCATTTGCTACTGATAGTGGCGGATGCCACATAAATTTGMAAAGCTGACTACTA 240
Db 1324 AAACCATTTGCTACTGATAGTGGCGGATGCCACATAAATTTGMAAAGCTGACTACTA 1383
QY 241 AAGGCTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAGGTC 300
Db 1384 AAGGCTATTCAAGAACAAATTGATCGCTAACGCTCCACAGTAACGACGACTACTTTGAGGTC 1443
QY 301 ATTGATTTTCAAGCGATGCAACCACTTACTGATCGAAACGGCAAGCTCTACTTTGCTGAC 360
Db 1444 ATTGATTTTCAAGCGATGCAACCACTTACTGATCGAAACGGCAAGCTCTACTTTGCTGAC 1503
QY 361 AAGATGTTTCGTTAACCTTTCGCGACCCCACTGCTCCAGAAATTTTGTTCAGCGGACAT 420
Db 1504 AAGATGTTTCGTTAACCTTTCGCGACCCCACTGCTCCAGAAATTTTGTTCAGCGGACAT 1563
QY 421 GTGCGGTTTAGACCATATATAAGAAAAAACAATCAAAATCAAGCGAAATCTGTTGATGTG 480
Db 1564 GTGCGGTTTAGACCATATATAAGAAAAAACAATCAAAATCAAGCGAAATCTGTTGATGTG 1623
QY 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 540
Db 1624 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 1683
QY 541 GATATCTAGCTATTGAARACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600
Db 1684 GATATCTAGCTATTGAARACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1743
QY 601 GCTCAAGCAAAAGCAATTTTAAACCAAAACCCAGGCTATACGATTTATGAACGTGAC 660
Db 1744 GCTCAAGCAAAAGCAATTTTAAACCAAAACCCAGGCTATACGATTTATGAACGTGAC 1803
QY 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 720
Db 1804 TCCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 1863
QY 721 TTCTACTTACCATGTCAAAATCGGAACAAAGCTTTATGAGATCAATAAAAAATCTGCTCTG 780
Db 1864 TTCTACTTACCATGTCAAAATCGGAACAAAGCTTTATGAGATCAATAAAAAATCTGCTCTG 1923
QY 781 AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 840
Db 1924 AATGAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 1983
QY 841 GAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAATACTGTTTACCAATCAATACGTT 900
Db 1984 GAAAGCCGTATGATCCCTTTGATGCGAGTCACTTGAATACTGTTTACCAATCAATACGTT 2043
QY 901 GATGTCACACCAAGAAATGCTTAAAAAGCGAGAGCTCTTAAACAGTAGCGAACCGTAAC 960
Db 2044 GATGTCACACCAAGAAATGCTTAAAAAGTAGAGAGCTCTTAAACAGTAGCGAACCGTAAC 2103
QY 961 TTAGACTTCAGAGATTTATACGATCTCTGTTAGAGGCTAAACTACTCTCAACAATCTC 1020
Db 2104 TTAGACTTCAGAGATTTATACGATCTCTGTTAGAGGCTAAACTACTCTCAACAATCTC 2163
QY 1021 GATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAAAAGTAGAGGATTAATCAGGATGAC 1080

Db 2164 GATGCTTTTGGTATTATGAGTACTATACCTTAACCTGGAAGGTAGAGGATAATCACGATGAC 2223
Qy 1081 ACCAACCGTATCATAAACCGTTTATATGCGCAAGCGACCCGGAAGGAGAGATGCTAGCTAT 1140
Db 2224 ACCAACCGTATCATAAACCGTTTATATGCGCAAGCGACCCGGAAGGAGAGATGCTAGCTAT 2283
Qy 1141 CATTTAGCCCTATGATAAAGATCGTTTATATGCGCAAGCGACCCGGAAGGAGAGATGCTAGCTAT 1200
Db 2284 CATTTAGCCCTATGATAAAGATCGTTTATATGCGCAAGCGACCCGGAAGGAGAGATGCTAGCTAT 2343
Qy 1201 CGTTATACAGGACACCTATACCTGATGATAACCTTAACGACAAA 1242
Db 2344 CGTTATACAGGACACCTATACCTGATGATAACCTTAACGACAAA 2385

RESULT 15
ABA05547
ID ABA05547 standard; DNA; 8893 BP.
AC
XX ABA05547;
XX
DT 26-FEB-2002 (first entry)
XX
XX Maxadilan-streptokinase fusion protein plasmid pTVB3maxstk.
DE Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;
XX vasodilator; thrombolytic; angina; myocardial infarction; stroke;
KW gene therapy; maxadilan; sand fly; plasmid; ds.
XX
OS Lutzomyia longipalpis.
OS Unidentified.
OS Synthetic.
OS Chimeric.
XX
XX WO200185100-A2.
XX
XX 15-NOV-2001.
XX
XX 10-MAY-2001; 2001WO-US015209.
XX
XX 11-MAY-2000; 2000US-00569920.
XX
XX (GEHO) GEN HOSPITAL CORP.
XX
XX Reddy VB, Lerner E;
XX
XX WPI; 2002-062184/08.
XX
XX New fusion protein or conjugate, useful for treating unstable angina,
PT acute myocardial infarction or stroke, comprises a vasodilator
PT polypeptide and a thrombolytic polypeptide, or active fragments of the
PT polypeptides.
XX
XX Example 1; Fig 3; 37pp; English.
PS
XX The invention relates to a fusion protein or a conjugate comprising a
CC vasodilator polypeptide, or its active fragment, and a thrombolytic
CC polypeptide or its active fragment. The protein is useful for treating a
CC subject suffering from a partially or totally occluded blood vessel,
CC causing unstable angina, acute myocardial infarction or stroke. The
CC polynucleotide encoding the polypeptide is useful in gene therapy. The
CC vasodilatory action of the protein allows for the use of lower doses of a
CC thrombolytic while maintaining the clot dissolving effectiveness of the
CC thrombolytic, and the use of lower doses of the thrombolytic reduces
CC associated side effects. The present sequence is the plasmid used for the
CC preparation of Maxadilan-Streptokinase fusion protein. It contains
CC maxadilan cDNA, which encodes a vasodilator peptide produced by the
CC salivary gland of the New World sand fly, and streptokinase cDNA
XX
SQ Sequence 8893 BP; 2366 A; 2106 C; 2246 G; 2175 T; 0 U; 0 Other;
Query Match 98.1%; Score 1221.2; DB 6; Length 8893;

Best Local Similarity 99.0%; Pred. No. 0;
Matches 1229; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 1 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATAGTTGTT 60
Db 5935 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATAGTTGTT 5994
Qy 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAAT 120
Db 5995 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTTTGAAT 6054
Qy 121 GACCTAATCATCAGACCTGCTCATGAGGAGAAACAGACAGCAAGGCTTAACTCCAAATCA 180
Db 6055 GATCTAATCATCAGACCTGCTCATGAGGAGAAACAGACAGCAAGGCTTAACTCCAAATCA 6114
Qy 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAAAAGCTGACTTACTA 240
Db 6115 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAAAAGCTGACTTACTA 6174
Qy 241 AAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAACGACGACTACTTTAGGTC 300
Db 6175 AAGGCTATTCAAGAAACAATTTGATCGCTAACGTCACAGTAACGACGACTACTTTAGGTC 6234
Qy 301 ATTGATTTTTCGAGCGATGCAACCATTTACTGATCGAAACGCGAAGGCTCTACTTTGCTGAC 360
Db 6235 ATTGATTTTTCGAGCGATGCAACCATTTACTGATCGAAACGCGAAGGCTCTACTTTGCTGAC 6294
Qy 361 AAAGATGGTTTCGGTAACTTTGCCGACCCCACTGTCCAAGAAATTTTGTCTAAGCGACAT 420
Db 6295 AAAGATGGTTTCGGTAACTTTGCCGACCCCACTGTCCAAGAAATTTTGTCTAAGCGACAT 6354
Qy 421 GTGCGGCTTAGACCATATAAAGAAAAACCAATACAAATCAAGGAAATCTGTTGATGTG 480
Db 6355 GTGCGGCTTAGACCATATAAAGAAAAACCAATACAAATCAAGGAAATCTGTTGATGTG 6414
Qy 481 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 540
Db 6415 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 6474
Qy 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 600
Db 6475 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTACTA 6534
Qy 601 GCTCAGACCAAAAGCATTTTAAACAAACCCAGGCTATACGATTTATGACGATGAC 660
Db 6535 GCTCAGACCAAAAGCATTTTAAACAAACCCAGGCTATACGATTTATGACGATGAC 6594
Qy 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 720
Db 6595 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 6654
Qy 721 TTTACTTTACCATGTCAAAAATTCGGGAAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
Db 6655 TTTACTTTACCATGTCAAAAATTCGGGAAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 6714
Qy 781 AATGAGAAATAAACACACTGACCTGATCTCTGAGGAAATATTACGTCCTTAAAAAAGGG 840
Db 6715 AATGAGAAATAAACACACTGACCTGATCTCTGAGGAAATATTACGTCCTTAAAAAAGGG 6774
Qy 841 GAAAAGCCGATGATCCCTTTGATCGCAGTCACCTTGAAAACCTGTTTCCACCATCAATACGTT 900
Db 6775 GAAAAGCCGATGATCCCTTTGATCGCAGTCACCTTGAAAACCTGTTTCCACCATCAATACGTT 6834
Qy 901 GATGTCACCAACCAAGATTTGCTTAAAGAGGAGAGAGCTCTTAAACAGTACGAGCAAGTAAAC 960
Db 6835 GATGTCACCAACCAAGATTTGCTTAAAGAGGAGAGAGCTCTTAAACAGTACGAGCAAGTAAAC 6894
Qy 961 TTAGACTTACAGATTTTATACGATCTCTGATGATGAGGCTAACTCTCTACAAATCTC 1020
Db 6895 TTAGACTTACAGATTTTATACGATCTCTGATGATGAGGCTAACTCTCTACAAATCTC 6954
Qy 1021 GATGCTTTTGGTATTATGAGACTATACCTTAACTGGAAAAAGTAGAGGATAATCACGATGAC 1080

Db	6955	GATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAAGTAGAGGATAATCACGATGAC	7014
Qy	1081	ACCAACCGTATCATTAACCGTTTATATGGCCAGCGACCCGAAAGGAGAGATGCTAGCTAT	1140
Db	7015	ACCAACCGTATCATTAACCGTTTATATGGCCAGCGACCCGAAAGGAGAGATGCTAGCTAT	7074
Qy	1141	CATTAGCCTATGATTAAGATCGTTTATACCGAAGAAAGACGAGAACTTTACAGCTACCTG	1200
Db	7075	CATTAGCCTATGATTAAGATCGTTTATACCGAAGAAAGACGAGAACTTTACAGCTACCTG	7134
Qy	1201	CGTTATACAGGACACCTTATACCTGATACCCCTAACGACAAA	1242
Db	7135	CGTTATACAGGACACCTTATACCTGATACCCCTAACGACAAA	7176

Search completed: May 6, 2004, 03:54:39
Job time : 555 secs

et.al.
 High level expression of streptokinase in *Escherichia coli*
 Biotechnology (N.Y.) 10 (10), 1138-1142 (1992)
 93000998
 MEDLINE
 1368792
 PUBMED
 GenBank staff at the National Library of Medicine created this
 entry [NCBI gisseq 115306] from the original journal article.
 This sequence comes from Fig. 2A.

FEATURES
 source
 1..1245
 /organism="Streptococcus dysgalactiae subsp. equisimilis"
 /mol_type="genomic DNA"
 /db_xref="taxon:119602"
 /note="group: C"
 1..1245
 /gene="SKC-2"
 1..1245
 /gene="SKC-2"
 /note="SKC-2"
 /note="streptokinase; This sequence comes from Fig. 2B;
 Mature protein lacks initial Met"
 /codon_start=1
 /product="SKC-2"
 /protein_id="AAC60418.1"
 /db_xref="GI:257197"
 /translation="IAGFELLDRPVSNNQSLVSVAGTVETGNNQDLSLKPFEDLTS
 RPAGGTGEGLSKPKFPFADSGMPHKLKDLKAIQELIANVHSDYFEVID
 FASDATITDRNGKVFADKGSVLTPTQVEFLLSGHVRVYKPKPIQONAKSVDV
 EYTVQTPFLPDDDFRPLKDTLLKLTLAGDTTSQELLAQASILNTHPGVTYVE
 RDSIVTHDNDIFRTILPDMQOEFYHVNREQAEYINKKSLNEEINNTDLSIKYIV
 LKGEKPYDPDRSHLKLFTIKVVDVNTNELLSEQLLTASERNLPDRILDYDRDAK
 LKYNLDAPGIMDYLTKGVEDNDDNRIITVYMKRPEGENNSHYLADKDYTEE
 BREVISYLRITGTPIDPNNDK"

ORIGIN

Query Match 100.0%; Score 1245; DB 1; Length 1245;
 Best Local Similarity 100.0%; Pred. No. 5.2e-297;
 Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACGCAATAGTTGTT 60
 Db 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACGCAATAGTTGTT 60

Qy 61 AGCGTTCCTGCTACTGTTGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATTT 120
 Db 61 AGCGTTCCTGCTACTGTTGGGGACGAATCAAGACATTAAGTCTTAATTTTGAATTT 120

Qy 121 GACCTAACATCAGCACCTGCTCATGGAGGAAAGACAGACGCAAGGCTTAAGTCCAAATCA 180
 Db 121 GACCTAACATCAGCACCTGCTCATGGAGGAAAGACAGACGCAAGGCTTAAGTCCAAATCA 180

Qy 181 AAACATTTGCTACTGATAGTGGCGGATGCGACATTAACCTGAAAGCTGACTTACTA 240
 Db 181 AAACATTTGCTACTGATAGTGGCGGATGCGACATTAACCTGAAAGCTGACTTACTA 240

Qy 241 AAGCTATTCAAGAACAAATGATCGCTAACTCCACAGTACGACGACTCTTTGAGGTC 300
 Db 241 AAGCTATTCAAGAACAAATGATCGCTAACTCCACAGTACGACGACTCTTTGAGGTC 300

Qy 301 ATTGATTTGCAAGCGATGCAACATTAAGTATGATGCAACGCAAGGCTTACTTTGCTGAC 360
 Db 301 ATTGATTTGCAAGCGATGCAACATTAAGTATGATGCAACGCAAGGCTTACTTTGCTGAC 360

Qy 361 AAAGATGGTTCGGTAACCTTGGCCGACCGACCTGTCAGAAATTTTGGCTAAGCGGCAT 420
 Db 361 AAAGATGGTTCGGTAACCTTGGCCGACCGACCTGTCAGAAATTTTGGCTAAGCGGCAT 420

Qy 421 GTGCGGTTAGACCATATAAGAAACCAATACAAATCAAGGAAATCTGTTGATGTG 480
 Db 421 GTGCGGTTAGACCATATAAGAAACCAATACAAATCAAGGAAATCTGTTGATGTG 480

Qy 481 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTACAGCAGGCTCTCAA 540
 Db 481 GAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTACAGCAGGCTCTCAA 540

Qy 541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600
 Db 541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600

Qy 601 GCTCAGACCAAGCAATTTTAAACAAACCCAGGCTATAGATTATGAACGTGAC 660
 Db 601 GCTCAGACCAAGCAATTTTAAACAAACCCAGGCTATAGATTATGAACGTGAC 660

Qy 661 TCCTCAATCTGCTACTCATGACAAATGCAATTTCCGTACGATTTTACCAATGGATCAAG 720
 Db 661 TCCTCAATCTGCTACTCATGACAAATGCAATTTCCGTACGATTTTACCAATGGATCAAG 720

Qy 721 TTTTACTTACCATGTCAAAAATCGGAAACAAAGCTTTATGAGATCAATAAAAAATCTGG 780
 Db 721 TTTTACTTACCATGTCAAAAATCGGAAACAAAGCTTTATGAGATCAATAAAAAATCTGG 780

Qy 781 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
 Db 781 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840

Qy 841 GAAAAGCCGTATCATCCCTTTGATCGCAGTCACCTTGAACCTGTTCCACATCAAAATAG 900
 Db 841 GAAAAGCCGTATCATCCCTTTGATCGCAGTCACCTTGAACCTGTTCCACATCAAAATAG 900

Qy 901 GATGTCAACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAAC 960
 Db 901 GATGTCAACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAAC 960

Qy 961 TTAGACTTCAGAGATTTATAGATCTCTCGTGAAGAGTAACTCTCTACCAACATCTC 1020
 Db 961 TTAGACTTCAGAGATTTATAGATCTCTCGTGAAGAGTAACTCTCTACCAACATCTC 1020

Qy 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1080
 Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGAGTAGAGGATAATCAGATGAC 1080

Qy 1081 ACCAAGCGTATCATACCGTTTATATGGGCAACGCGACCCGAAAGAGAGATGCTAGCTAT 1140
 Db 1081 ACCAAGCGTATCATACCGTTTATATGGGCAACGCGACCCGAAAGAGAGATGCTAGCTAT 1140

Qy 1141 CATTTAGCTATCATTAAGATCGTTATACGGAAGAGAGAGATGTTACAGCTACCTG 1200
 Db 1141 CATTTAGCTATCATTAAGATCGTTATACGGAAGAGAGAGATGTTACAGCTACCTG 1200

Qy 1201 CGTTATACGGGACACCTATACCTGATTAACCTTAACGACAAATAA 1245
 Db 1201 CGTTATACGGGACACCTATACCTGATTAACCTTAACGACAAATAA 1245

RESULT 2

A20015
 LOCUS
 DEFINITION
 SEQ ID NO: 23; Nucleotide sequence for methionol-streptokinase fusion protein.
 ACCESSION
 A20015
 VERSION
 A20015.1
 KEYWORDS
 A20015.1
 ORGANISM
 A20015.1
 G1:1247848
 SYNTHETIC CONSTRUCT
 SYNTHETIC CONSTRUCT
 ARTIFICIAL SEQUENCES
 1 (bases 1 to 1257)
 PROTEINS AND NUCLEIC ACIDS
 Patent: WO 9109125-A 23 27-JUN-1991;
 Location/Qualifiers
 1..1257
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 4..1251
 /codon_start=1
 /transl_table=11

CDS

```

/product="methionyl-streptokinase fusion protein"
/protein_id="CAA01486.1"
/db_xref="GI:1247849"
/db_xref="REMRMBL:CAA01486"
/translation="MIAGPEWLDRSPVNNNSQLVSVAGTVEGTNODISLKFIEDLT
SRPAHGKTEQGLSPKSPKFPATDSGAMPKLEKADLLKAIQELIANVHSNDYFPIV
DFASDATIIDRNGKVPFADKQDSVTLPTQVQFLLSGHVRVPYKEKPIQNAQKSV
VEYVQFTPLNPDPRGLKDTKLKTLAIGDTTSOELLAAQOSILNKTHPGYTIY
ERDSSITVTHNDI.FRIILPMDOSEFTVHKRBOAYEINKKSGLINEINNDLISEKYV
VLKKGKPYDPDRSHLKLFTIKYDVYNTNELLKSEQLLTASERNLDPDRDLDPDKA
KLINLDAFGIMDYTLTKRVEDNDDTNRILITVYMKRPEGENASYHLAYDKORYTE
EBREVSXYLRYTGTPTPDNPNDK"

Query Match          99.7%;   Score 1241.8;   DB 6;   Length 1257;
Best Local Similarity 99.8%;   Pred. No. 3.2e-296;
Matches 1243;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

QY      1  ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTAGTTGTT 60
DB      7  ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTAGTTGTT 66

QY      61  AGCGTTGCTGCTACTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATTT 120
DB      67  AGCGTTGCTGCTACTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATTT 126

QY      121  GACCTAACATCAGACCTGTCTATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 180
DB      127  GACCTAACATCAGACCTGTCTATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 186

QY      181  AAACCATTTGCTACTGATAGTGGCGGATGCCATATAAATTGTAAGTCCAAATCA 240
DB      187  AAACCATTTGCTACTGATAGTGGCGGATGCCATATAAATTGTAAGTCCAAATCA 246

QY      241  AAGGCTATTCAAGAACAAATTGATCCCTACGTCCTCAAGTAAACGACCTACTTGGAGTTC 300
DB      247  AAGGCTATTCAAGAACAAATTGATCCCTACGTCCTCAAGTAAACGACCTACTTGGAGTTC 306

QY      301  ATTGATTTTGCAGCGCATGCAACATTACTGTGATGCGAAACGGCAAGCTCTACTTTGCTGAC 360
DB      307  ATTGATTTTGCAGCGCATGCAACATTACTGTGATGCGAAACGGCAAGCTCTACTTTGCTGAC 366

QY      361  AAGATGTTTGGTAACCTTGGCGACCCCACTGTGTCAGAGAAATTTTGTGTAAGCGACAT 420
DB      367  AAGATGTTTGGTAACCTTGGCGACCCCACTGTGTCAGAGAAATTTTGTGTAAGCGACAT 426

QY      421  GTGCGGTTAGACCATATAAGAGAAACCAATACAAATCAAGCGAAATCTGTGTGATGTG 480
DB      427  GTGCGGTTAGACCATATAAGAGAAACCAATACAAATCAAGCGAAATCTGTGTGATGTG 486

QY      481  GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGACAGCTCTCAA 540
DB      487  GAATATACCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTCAAGACAGCTCTCAA 546

QY      541  GATACCTAAGCTATTGAACAACCTAGTATCGGTGACACCATCACTCTCAAGAAATTA 600
DB      547  GATACCTAAGCTATTGAACAACCTAGTATCGGTGACACCATCACTCTCAAGAAATTA 606

QY      601  GCTCAAGCAACAAAGCATTTTAAACAAACCCAGGCTATACGATTTTATGAACCTGAC 660
DB      607  GCTCAAGCAACAAAGCATTTTAAACAAACCCAGGCTATACGATTTTATGAACCTGAC 666

QY      661  TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
DB      667  TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 726

QY      721  TTTACTTACCATGTCAAAATTCGGGAACCAAGCTTATGAGATCAATAAAAATCTGCTGT 780
DB      727  TTTACTTACCATGTCAAAATTCGGGAACCAAGCTTATGAGATCAATAAAAATCTGCTGT 786

QY      781  AATGAAGAATAAACCACTGACCTGATCTCTGAGAAATATTACCTCTTAAAAAGGG 840
DB      787  AATGAAGAATAAACCACTGACCTGATCTCTGAGAAATATTACCTCTTAAAAAGGG 846

```

```

QY      841  GAAAGCCGTATGATCCCTTTTGTATCGAGTCACTTGAATGTTTCAACATCAATAGCTT 900
DB      847  GAAAGCCGTATGATCCCTTTTGTATCGAGTCACTTGAATGTTTCAACATCAATAGCTT 906

QY      901  GATGTCAACACCAACCAATTTGCTTAAAGCGAGCAGCTCTTAAACAGCTACGCAACGTAAC 960
DB      907  GATGTCAACACCAACCAATTTGCTTAAAGCGAGCAGCTCTTAAACAGCTACGCAACGTAAC 966

QY      961  TTAGACTTCAGAGATTATACGATCTCTGTGATAGGCTTAACTACTCTTCAACAATCTTC 1020
DB      967  TTAGACTTCAGAGATTATACGATCTCTGTGATAGGCTTAACTACTCTTCAACAATCTTC 1026

QY      1021  GATGCTTTTGTATTTATGACTATATCTTAACTTGAAGAGTAGAGGATAATCAGATGAC 1080
DB      1027  GATGCTTTTGTATTTATGACTATATCTTAACTTGAAGAGTAGAGGATAATCAGATGAC 1086

QY      1081  ACCAACCGTATCATACCGTTTATATGGCAAGCCGCAAGGAGAGAGAAATGCTAGCTAT 1140
DB      1087  ACCAACCGTATCATACCGTTTATATGGCAAGCCGCAAGGAGAGAAATGCTAGCTAT 1146

QY      1141  CATTTAGCCTATGATTAAGATCGTTATACGGAAGAGAGAGAGAGTTTACAGCTACCTG 1200
DB      1147  CATTTAGCCTATGATTAAGATCGTTATACGGAAGAGAGAGAGAGTTTACAGCTACCTG 1206

QY      1201  CGTTATACAGGACACCTATACCTGTATAACCCCTTAAACGACAAATAA 1245
DB      1207  CGTTATACAGGACACCTATACCTGTATAACCCCTTAAACGACAAATAA 1251

RESULT 3
I13203
LOCUS      I13203              1257 bp      DNA      linear      PAT 26-JUL-1995
DEFINITION Sequence 25 from patent US 5434073.
ACCESSION I13203
VERSION   I13203.1  GI:910551
KEYWORDS
SOURCE    Unknown.
ORGANISM  Unknown.
REFERENCE Unclassified.
AUTHORS   Dawson,K., Hunter,M.G. and Czaplewski,L.G.
TITLE     Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL   Patent: US 5434073-A 25 18-JUL-1995;
FEATURES  Location/Qualifiers
           source
           1..1257
           /organism="unknown"
           /mol_type="unassigned DNA"

ORIGIN

Query Match          99.7%;   Score 1241.8;   DB 6;   Length 1257;
Best Local Similarity 99.8%;   Pred. No. 3.2e-296;
Matches 1243;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

QY      1  ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTAGTTGTT 60
DB      7  ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTAGTTGTT 66

QY      61  AGCGTTGCTGCTACTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATTT 120
DB      67  AGCGTTGCTGCTACTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATTT 126

QY      121  GACCTAACATCAGACCTGTCTATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 180
DB      127  GACCTAACATCAGACCTGTCTATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 186

QY      181  AAACCATTTGCTACTGATAGTGGCGGATGCCATATAAATTGTAAGTCCAAATCA 240
DB      187  AAACCATTTGCTACTGATAGTGGCGGATGCCATATAAATTGTAAGTCCAAATCA 246

QY      241  AAGGCTATTCAAGAACAAATTGATCCCTACGTCCTCAAGTAAACGACCTACTTGGAGTTC 300
DB      247  AAGGCTATTCAAGAACAAATTGATCCCTACGTCCTCAAGTAAACGACCTACTTGGAGTTC 306

```

QY 301 ATTGATTTTGAAGCGATGCAACCACTTACTGATGAAACGGCAAGGTCTACTTTGCTGAC 360
DB |||||
QY 307 ATTGATTTTGAAGCGATGCAACCACTTACTGATGAAACGGCAAGGTCTACTTTGCTGAC 366
DB |||||
QY 361 AAAAGATGTTCCGTTAACTCTCCGACCCCAACTGCTGCAAGAAATTTTGTCTAAGCGGCAT 420
DB |||||
QY 367 AAAAGATGTTCCGTTAACTCTCCGACCCCAACTGCTGCAAGAAATTTTGTCTAAGCGGCAT 426
DB |||||
QY 421 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAATCAAGGAATCTGTTGATGTG 480
DB |||||
QY 427 GTGCGCGTTAGACCATATAAAGAAAAACCAATACAAATCAAGGAATCTGTTGATGTG 486
DB |||||
QY 481 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACGATCTCAAA 540
DB |||||
QY 487 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGACGATCTCAAA 546
DB |||||
QY 541 GATATCTAAGCTATTGAAACCACTAGCTATCGGTGACCACTACATCTCAAGAAATTA 600
DB |||||
QY 547 GATATCTAAGCTATTGAAACCACTAGCTATCGGTGACCACTACATCTCAAGAAATTA 606
DB |||||
QY 601 GCTCAAGCACAAAGCATTTTAAACAAACCCAGCCAGCTATACGATTTATGAACGTGAC 660
DB |||||
QY 607 GCTCAAGCACAAAGCATTTTAAACAAACCCAGCCAGCTATACGATTTATGAACGTGAC 666
DB |||||
QY 661 TCCTCAATCGTCACTCATGACATGACATTTTCCGTAGCTTTTACCAATGGATCAAGAG 720
DB |||||
QY 667 TCCTCAATCGTCACTCATGACATGACATTTTCCGTAGCTTTTACCAATGGATCAAGAG 726
DB |||||
QY 721 TTTTACTTACCATGTCMAAAATCGGGACCAAGCTTATGAGATCAATAAAATCTGGTCTG 780
DB |||||
QY 727 TTTTACTTACCATGTCMAAAATCGGGACCAAGCTTATGAGATCAATAAAATCTGGTCTG 786
DB |||||
QY 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 840
DB |||||
QY 787 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 846
DB |||||
QY 841 GAAAGCGGTATGATCCCTTTGATGCGAGTCACTTTGAAATCTGTTCAACATCAAAATAGCT 900
DB |||||
QY 847 GAAAGCGGTATGATCCCTTTGATGCGAGTCACTTTGAAATCTGTTCAACATCAAAATAGCT 906
DB |||||
QY 901 GATGTCAACACCAAGCATTTCTAAAGCGGAGGAGCTTTAAACAGCTAGCGACGTAAC 960
DB |||||
QY 907 GATGTCAACACCAAGCATTTCTAAAGCGGAGGAGCTTTAAACAGCTAGCGACGTAAC 966
DB |||||
QY 961 TTAGACTTTCAGAGATTTATACGATCTCTGATGAGGCTAAACTACTCTACAAATCTC 1020
DB |||||
QY 967 TTAGACTTTCAGAGATTTATACGATCTCTGATGAGGCTAAACTACTCTACAAATCTC 1026
DB |||||
QY 1021 GATGCTTTTGGTATTTATGGAATATACCTTAACTGGAAGATAGAGGATAATCAAGTAC 1080
DB |||||
QY 1027 GATGCTTTTGGTATTTATGGAATATACCTTAACTGGAAGATAGAGGATAATCAAGTAC 1086
DB |||||
QY 1081 ACCAACCGTATCATAAACCGTTTATATGCGAAGGAGCCGAGGAGAGAGATGCTAGCTAT 1140
DB |||||
QY 1087 ACCAACCGTATCATAAACCGTTTATATGCGAAGGAGCCGAGGAGAGAGATGCTAGCTAT 1146
DB |||||
QY 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAACGAGAGTTTACAGCTACCTG 1200
DB |||||
QY 1147 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAACGAGAGTTTACAGCTACCTG 1206
DB |||||
QY 1201 CGTTATACAGGACCACTTATACCTGATTAACCTTACGCAATTA 1245
DB |||||
QY 1207 CGTTATACAGGACCACTTATACCTGATTAACCTTACGCAATTA 1251
DB |||||

RESULT 4

A20009 LOCUS A20009 1317 bp DNA linear PAT 14-JUL-1995
DEFINITION SEQ ID NO: 17; OmpA fused to a mature streptokinase gene.
ACCESSION A20009
VERSION A20009.1 GI:1247841
KEYWORDS

SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1317)
AUTHORS
TITLE PROTEINS AND NUCLEIC ACIDS
JOURNAL Patent: WO 9109125-A 17 27-JUN-1991;
FEATURES Location/Qualifiers
1..1317
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
4..1311
/note="unnamed protein product; ompA fused to
streptokinase gene"
/codon_start=1
/transl_table=11
/protein_id="CAA01485.1"
/db_xref="GI:1247842"
/db_xref="RENTREMBL:CAA01485"
translation="MKTAIAIAVALAGPATVQAIAAGPEWLLDRPVSNNLSQVSV
GTVEGNDIISLAFEDLTSRPAHSGKTEQGLSPKSPKPTDGSAMPHKLEADLK
AIGBQLIANVHNDIDFEVIDFASDATITDRNKVYFADKXGVTLPQVQFLLSG
HVRVPYKPKPIQNAQKSDVEYTVQTPINDIDFRLPMDQFTYHVKRQAYEN
ELLAQAQSLINKTHPGYTYIYERDSSIIVTHNDIFRILPMDQFTYHVKRQAYEN
KKSGLBEINNTDLISKKYVKKPKYDFDRSHLKLFTIKYVDVNTVELLSQ
LTASERNLDPRLYDPRDKAKLLNNLDAGIMDYTLTGKVDNHDNRIITVYMGK
RPGENASYHLAYDKORYTEEREVSYLAYTGTPIPDNPDK"

CDS

1 ATTGCTCGACTGAGTGGCTGAGACCGTCCATCTCTCAACACAGCAATAGTTGTT 60
67 ATTGCTCGACTGAGTGGCTGAGACCGTCCATCTCTCAACACAGCAATAGTTGTT 126
61 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAAT 120
127 AGCGTTCTGCTACTGTTGAGGGGACGAATCAAGACATAGTCTTAAATTTTGAAT 186
121 GACCTAATCATCAGACCTGCTCATGAGGAGAAAGACAGACGAGCTTAAGTCCAAATCA 180
187 GACCTAATCATCAGACCTGCTCATGAGGAGAAAGACAGACGAGCTTAAGTCCAAATCA 246
181 AAACCATTTGCTACTGATGAGCGGATGCGCATATAAATTGAAAAGCTGACTACTA 240
247 AAACCATTTGCTACTGATGAGCGGATGCGCATATAAATTGAAAAGCTGACTACTA 306
241 AAGCTTATCAAGAACAAATGATGCTTAAGTCCACAGTAACGACGACTACTTTGAGTC 300
307 AAGCTTATCAAGAACAAATGATGCTTAAGTCCACAGTAACGACGACTACTTTGAGTC 366
301 ATTGATTTTCAAGCGATGCAACCATTTACTGATGAGAAACGGCAAGGTCTACTTTGCTGAC 360
367 ATTGATTTTCAAGCGATGCAACCATTTACTGATGAGAAACGGCAAGGTCTACTTTGCTGAC 426
361 AAAGATGTTTGGTAACTCTGCGACCCCACTGTGCAAGAAATTTTGTCTAAGCGGCAT 420
427 AAAGATGTTTGGTAACTCTGCGACCCCACTGTGCAAGAAATTTTGTCTAAGCGGCAT 486
421 GTGCGCGTTAGACCATATAAAGAAAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480
487 GTGCGCGTTAGACCATATAAAGAAAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 546
481 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGAGATTTTACAGACGATCTCAAA 540
547 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGAGATTTTACAGACGATCTCAAA 606
541 GATATCTAAGCTATTGAAACCACTAGCTATCGGTGACCACTACATCTCAAGAAATTA 600
607 GATATCTAAGCTATTGAAACCACTAGCTATCGGTGACCACTACATCTCAAGAAATTA 666

ORIGIN

Query Match 99.7%; Score 1241.8; DB 6; Length 1317;
Best Local Similarity 99.8%; Pred. No. 3.2e-236;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 601 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 660
DB 667 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 726
QY 661 TCTCAATCGTCTCATGACAAATGACATTTTCCGTCGATGATTTTACAAATGGATCAAGAG 720
DB 727 TCTCAATCGTCTCATGACAAATGACATTTTCCGTCGATGATTTTACAAATGGATCAAGAG 786
QY 721 TTTTACTTACCATCTCAAAATCGGAAACGAGCTTATGAGATCAATAAATAATCGTCTG 780
DB 787 TTTTACTTACCATCTCAAAATCGGAAACGAGCTTATGAGATCAATAAATAATCGTCTG 846
QY 781 AATGAAGAAATAAACCAACTGACCTGTCTCTGAGAAATATTTACGTCCTTAAAAAGGG 840
DB 847 AATGAAGAAATAAACCAACTGACCTGTCTCTGAGAAATATTTACGTCCTTAAAAAGGG 906
QY 841 GAAAGCCGTATGATCCCTTTGATCCGAGTCACTTGAATTAATGATGATCAATCAATACGTT 900
DB 907 GAAAGCCGTATGATCCCTTTGATCCGAGTCACTTGAATTAATGATGATCAATCAATACGTT 966
QY 901 GATGTCACCAACCAAGATTTGCTTAAAGCGAGCAGCTCTTAAACAGCTAGCGAAGTAA 960
DB 967 GATGTCACCAACCAAGATTTGCTTAAAGCGAGCAGCTCTTAAACAGCTAGCGAAGTAA 1026
QY 961 TTAGACTTCAGAGATTTATACGATCCCTGTGATAGGCTAAACTACTCTACAAATCTC 1020
DB 1027 TTAGACTTCAGAGATTTATACGATCCCTGTGATAGGCTAAACTACTCTACAAATCTC 1086
QY 1021 GATGCTTTGGTATTTATGAGCTATACCTTAACTGAAATCTGTTTCAACATCAATACGTT 1080
DB 1087 GATGCTTTGGTATTTATGAGCTATACCTTAACTGAAATCTGTTTCAACATCAATACGTT 1146
QY 1081 ACCAAGCGTATCATACCGTTTATATGCGAAGCGAGCAGCTCTTAAACAGCTAGCGAAGTAA 1140
DB 1147 ACCAAGCGTATCATACCGTTTATATGCGAAGCGAGCAGCTCTTAAACAGCTAGCGAAGTAA 1206
QY 1141 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAAACGAGAGTGTTCAGCTACCTG 1200
DB 1207 CATTTAGCCTATGATAAAGATCGTTTATACCGAAGAAACGAGAGTGTTCAGCTACCTG 1266
QY 1201 CTTTATACGGACACCTATACCTGATTAACCTTAAACAGCAATAA 1245
DB 1267 CTTTATACGGACACCTATACCTGATTAACCTTAAACAGCAATAA 1311
RESULT 5
LOCUS I13197 1317 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 18 from patent US 5434073.
ACCESSION I13197
VERSION I13197.1 GI:910545
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
1 (bases 1 to 1317)
AUTHORS Dawson,K., Hunter,M.G. and Czaplowski,L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 18 JUL-1995;
FEATURES
Location/Qualifiers
source 1. .1317
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 99.7%; Score 1241.8; DB 6; Length 1317;
Best Local Similarity 99.8%; Pred. No. 3.2e-296;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAAGCAATTTAGTTGT 60
DB 67 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACAAAGCAATTTAGTTGT 126
```

```
QY 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTTGAAATT 120
DB 127 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTTGAAATT 186
QY 121 GACCTAAACATCAACGACCTCTCATGAGGAAAGACAGACGAAGGCTTAAAGTCCAAATCA 180
DB 187 GACCTAAACATCAACGACCTCTCATGAGGAAAGACAGACGAAGGCTTAAAGTCCAAATCA 246
QY 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATATAAATTGAAAGCTGACTTACTA 240
DB 247 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATATAAATTGAAAGCTGACTTACTA 306
QY 241 AAGGCTTATTCAAGAACCAATTGATCGTAACTGCTCAAGTAAAGCAAGGCTTAAAGTCC 300
DB 307 AAGGCTTATTCAAGAACCAATTGATCGTAACTGCTCAAGTAAAGCAAGGCTTAAAGTCC 366
QY 301 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGCGAAGGCTTAAAGTCCGAC 360
DB 367 ATTGATTTTGAAGCGATGCAACCAATTACTGATCGAAACGCGAAGGCTTAAAGTCCGAC 426
QY 361 AAGGATGGTTCGGTAACTTGGCGACCCCAACCTGTCGAAGAAATTTTGTCTAAGCGGACAT 420
DB 427 AAGGATGGTTCGGTAACTTGGCGACCCCAACCTGTCGAAGAAATTTTGTCTAAGCGGACAT 486
QY 421 GTGCGCGTTAGACCATATAAAGAAAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
DB 487 GTGCGCGTTAGACCATATAAAGAAAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 546
QY 481 GAATATCTGTCAGTCTTACTCCCTTAAACCCCTGATGAGATTTAGACAGGCTCTCAAA 540
DB 547 GAATATCTGTCAGTCTTACTCCCTTAAACCCCTGATGAGATTTAGACAGGCTCTCAAA 606
QY 541 GATATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 600
DB 607 GATATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 666
QY 601 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 660
DB 667 GCTCAAGCAAAAGCAATTTTAAACAAAACCCAGGCTATACGATTTTATGAACGTGAC 726
QY 661 TCTTCAATCGTCACTCATGCAATGACATTTTCGGTACGATTTTACCAGTATGATCAAGAG 720
DB 727 TCTTCAATCGTCACTCATGCAATGACATTTTCGGTACGATTTTACCAGTATGATCAAGAG 786
QY 721 TTTTACTTACCATGTCAAAATCGGAAACGAGCTTATGAGATCAATAAATAATCTGCTCTG 780
DB 787 TTTTACTTACCATGTCAAAATCGGAAACGAGCTTATGAGATCAATAAATAATCTGCTCTG 846
QY 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 840
DB 847 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTTACGTCCTTAAAAAGGG 906
QY 841 GAAAGCCGTATGATCCCTTTGATCCGAGTCACTTGAATTAATGATGATCAATCAATACGTT 900
DB 907 GAAAGCCGTATGATCCCTTTGATCCGAGTCACTTGAATTAATGATGATCAATCAATACGTT 966
QY 901 GATGTCACCAACCAAGATTTGCTTAAAGCGAGCAGCTCTTAAACAGCTAGCGAAGTAA 960
DB 967 GATGTCACCAACCAAGATTTGCTTAAAGCGAGCAGCTCTTAAACAGCTAGCGAAGTAA 1026
QY 961 TTAGACTTCAGAGATTTATACGATCCCTGTGATAGGCTAAACTACTCTACAAATCTC 1020
DB 1027 TTAGACTTCAGAGATTTATACGATCCCTGTGATAGGCTAAACTACTCTACAAATCTC 1086
QY 1021 GATGCTTTGGTATTTATGAGCTATACCTTAACTGAAATCTGTTTCAACATCAATACGTT 1080
DB 1087 GATGCTTTGGTATTTATGAGCTATACCTTAACTGAAATCTGTTTCAACATCAATACGTT 1146
QY 1081 ACCAAGCGTATCATACCGTTTATATGCGAAGCGACCCGAAGGAGAGAAATCTAGCTAT 1140
DB 1147 ACCAAGCGTATCATACCGTTTATATGCGAAGCGACCCGAAGGAGAGAAATCTAGCTAT 1206
```


REFERENCE 1 (bases 1 to 1335)
AUTHORS Dawson,K., Hunter,M.G. and Czaplewski,L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 14 18-JUL-1995;
FEATURES Location/Qualifiers
source 1..1335
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 99.7%; Score 1241.8; DB 6; Length 1335;
Best Local Similarity 99.8%; Pred. No. 3.2e-296;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATTAGTTGTT 60
Db 85 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATTAGTTGTT 144
QY 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATT 120
Db 145 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAATT 204
QY 121 GACCTAACATCAGACCTGCTCATGAGAGGAAAGACAGCAGCAAGCTTAAAGTCCAAATCA 180
Db 205 GACCTAACATCAGACCTGCTCATGAGAGGAAAGACAGCAGCAAGCTTAAAGTCCAAATCA 264
QY 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCATATAAACTTGAAAAGCTGACTTACTA 240
Db 265 AAACCAATTTGCTACTGATAGTGGCGGATGCCATATAAACTTGAAAAGCTGACTTACTA 324
QY 241 AAGGCTATTCAAGACAAATTGATCGCTTAACGTCCACAGTAAACGCACTACTTTGAGGTC 300
Db 325 AAGGCTATTCAAGACAAATTGATCGCTTAACGTCCACAGTAAACGCACTACTTTGAGGTC 384
QY 301 ATTGATTTTCCAGCGATGCAACCATTTACTGATGCGAAACGCAAGGTCTACTTTGCTGAC 360
Db 385 ATTGATTTTCCAGCGATGCAACCATTTACTGATGCGAAACGCAAGGTCTACTTTGCTGAC 444
QY 361 AAGATGTTTGGTAACTTGGCCGACCAACCTGTCCAGAAATTTTGTCTAAAGCGACAT 420
Db 445 AAGATGTTTGGTAACTTGGCCGACCAACCTGTCCAGAAATTTTGTCTAAAGCGACAT 504
QY 421 GTGCGGTTAGACCATATAAGAAACCAATACAAATCAAAATCAAGCGAAATCTGTTGATGTG 480
Db 505 GTGCGGTTAGACCATATAAGAAACCAATACAAATCAAAATCAAGCGAAATCTGTTGATGTG 564
QY 481 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAAGCAGGTCTCAAA 540
Db 565 GAATATAGTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTCAAGCAGGTCTCAAA 624
QY 541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 600
Db 625 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTAATA 684
QY 601 GCTCAAGCACAAGCAATTTTAAACAAACCCAGGCTATACGATTTTATGAACTGAC 660
Db 685 GCTCAAGCACAAGCAATTTTAAACAAACCCAGGCTATACGATTTTATGAACTGAC 744
QY 661 TCCTCAATCGTCACTCATGCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 720
Db 745 TCCTCAATCGTCACTCATGCAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 804
QY 721 TTTTACTTACATGTCAAAATTCGGGAACCAAGCTTATGAGATCAATAAAAATCTGCTCTG 780
Db 805 TTTTACTTACATGTCAAAATTCGGGAACCAAGCTTATGAGATCAATAAAAATCTGCTCTG 864
QY 781 AATGAGAAATAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 840
Db 865 AATGAGAAATAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAGGG 924
QY 841 GAAAAGCCGATGATCCCTTTTGTATCGCAGTCACCTTGAAACTGTTCCACATCAATACGTT 900
Db 925 GAAAAGCCGATGATCCCTTTTGTATCGCAGTCACCTTGAAACTGTTCCACATCAATACGTT 984

QY 901 GATGTCAACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGACGTAAC 960
Db 985 GATGTCAACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGACGTAAC 1044
QY 961 TTAGACTTCAGAGATTTTATACGATTCCTCGTGAAGGCTAAACTTACTCTACAAATCTC 1020
Db 1045 TTAGACTTCAGAGATTTTATACGATTCCTCGTGAAGGCTAAACTTACTCTACAAATCTC 1104
QY 1021 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATATCAGATGAC 1080
Db 1105 GATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGAGTAGAGGATATCAGATGAC 1164
QY 1081 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
Db 1165 ACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1224
QY 1141 CATTTAGCTTATGATTAAGATTCCTTATACCGAAGAGAACGAGAGTTTACAGTACCTG 1200
Db 1225 CATTTAGCTTATGATTAAGATTCCTTATACCGAAGAGAACGAGAGTTTACAGTACCTG 1284
QY 1201 CGTTATACAGGACACCTATACCTGATACCCCTAACGACAAATAA 1245
Db 1285 CGTTATACAGGACACCTATACCTGATACCCCTAACGACAAATAA 1329

RESULT 8
A20027
LOCUS A20027 1458 bp DNA linear PAT 14-JUL-1995
DEFINITION SEQ ID NO: 35; Synthetic nucleotide sequence for Hirudin-streptokinase fusion protein.
ACCESSION A20027
VERSION A20027.1 GI:1247862
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 1458)
AUTHORS PROTEINS AND NUCLEIC ACIDS
TITLE Patent: WO 9109125-A 35 27-JUN-1991;
JOURNAL /organism="synthetic construct"
FEATURES /mol_type="unassigned DNA"
Source /db_xref="taxon:32630"
Location/Qualifiers
1..1458
ORIGIN
Query Match 99.7%; Score 1241.8; DB 6; Length 1458;
Best Local Similarity 99.8%; Pred. No. 3.2e-296;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATTAGTTGTT 60
Db 208 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACAGCCCAATTAGTTGTT 267
QY 61 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGCATTAGTCTTAAATTTTGGAAATT 120
Db 268 AGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGCATTAGTCTTAAATTTTGGAAATT 327
QY 121 GACCTAACATCAGACCTGCTCATGAGGAGAAAGACAGACGAGGCTTAAAGTCCAAATCA 180
Db 328 GACCTAACATCAGACCTGCTCATGAGGAGAAAGACAGACGAGGCTTAAAGTCCAAATCA 387
QY 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCATATAAACTTGAAAAGCTGACTTACTA 240
Db 388 AAACCAATTTGCTACTGATAGTGGCGGATGCCATATAAACTTGAAAAGCTGACTTACTA 447
QY 241 AAGGCTATTCAAGACAAATTTGATCGCTTAAAGTAAAGAGTAGAGGATATCAGTACCTG 300
Db 448 AAGGCTATTCAAGACAAATTTGATCGCTTAAAGTAAAGAGTAGAGGATATCAGTACCTG 507
QY 301 ATTGATTTTGAAGCGATGCAACCATTTACTGATGCGAAACGCGCAAGGTCTACTTTGCTGAC 360

Db 508 ATTGATTTTGCAGGATGCAACCACTTACTGATCCAAACGCGAAGGTCTACTTTTCTGCAC 567
Qy 361 AAAAGATGTTTGGTAACTTGCAGCCCAACCTGTCGAAGAAATTTTGTCTAGCGGACAT 420
Db 568 AAAAGATGTTTGGTAACTTGCAGCCCAACCTGTCGAAGAAATTTTGTCTAGCGGACAT 627
Qy 421 GTGCGCGTTAGACCATATAAAGAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
Db 628 GTGCGCGTTAGACCATATAAAGAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 687
Qy 481 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCCAGGTCTCAAA 540
Db 688 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCCAGGTCTCAAA 747
Qy 541 GATCTAAGCTATTGAAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 600
Db 748 GATCTAAGCTATTGAAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 807
Qy 601 GCTCAAGCAACAAGCAATTTTAAACAAACCCACCCAGGCTATACGATTTATGAACGTGAC 660
Db 808 GCTCAAGCAACAAGCAATTTTAAACAAACCCACCCAGGCTATACGATTTATGAACGTGAC 867
Qy 661 TCCTCAATCTGTCACATCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 720
Db 868 TCCTCAATCTGTCACATCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 927
Qy 721 TTTACTTACCATTGTCGGAACCAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
Db 928 TTTACTTACCATTGTCGGAACCAAGCTTATGAGATCAATAAAAAATCTGGTCTG 987
Qy 781 AATGAGAAATAAACAACACCTGATCTCTGAGAAATATAGCTCTTAAAAAGGG 840
Db 988 AATGAGAAATAAACAACACCTGATCTCTGAGAAATATAGCTCTTAAAAAGGG 1047
Qy 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 900
Db 1048 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 1107
Qy 901 GATGTCACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAAGTAA 960
Db 1108 GATGTCACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAAGTAA 1167
Qy 961 TTAGATTTTCAAGATTTATACGATCTCTGATAGGCTTAACTCTACAAATCTC 1020
Db 1168 TTAGATTTTCAAGATTTATACGATCTCTGATAGGCTTAACTCTACAAATCTC 1227
Qy 1021 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTAGAGATAATCACGATGAC 1080
Db 1228 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGTAGAGATAATCACGATGAC 1287
Qy 1081 ACCAACCGTATCATNAACCGTTTATATGGAAGAGCCGACCCGAGGAGAGAAATGCTAGCTAT 1140
Db 1288 ACCAACCGTATCATNAACCGTTTATATGGAAGAGCCGACCCGAGGAGAGAAATGCTAGCTAT 1347
Qy 1141 CATTTAGCCTATGATAGAGATGTTTATACCGAGAGAACGAGAGTTTACAGTACCTG 1200
Db 1348 CATTTAGCCTATGATAGAGATGTTTATACCGAGAGAACGAGAGTTTACAGTACCTG 1407
Qy 1201 CGTTATACAGGACACCTATACCTGATACCTTAAACCTTAAACCAATAA 1245
Db 1408 CGTTATACAGGACACCTATACCTGATACCTTAAACCTTAAACCAATAA 1452

RESULT 9
LOCUS I13215 1458 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 42 from patent US 5434073.
ACCESSION I13215
VERSION I13215.1 GI:910563
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1458)
AUTHORS Dawson,K., Hunter,M.G. and Czaplowski,L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 42 18-JUL-1995;
FEATURES Location/Qualifiers
source 1..1458
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 99.7%; Score 1241.8; DB 6; Length 1458;
Best Local Similarity 99.8%; Pred. No. 3.2e-296;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATTGCTGACCTGTAGCGTCTGCTAGACGTCATCTGTCAACACAGCAATTAAGTTGTT 60
Db 208 ATTGCTGACCTGTAGCGTCTGCTAGACGTCATCTGTCAACACAGCAATTAAGTTGTT 267
Qy 61 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTTTAAATTTTGAAT 120
Db 268 AGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAAGTTTAAATTTTGAAT 327
Qy 121 GACCTAATCACAACCTGCTCATGAGGAGAAAGACAGAGCGTTAAGTCCAAATCA 180
Db 328 GACCTAATCACAACCTGCTCATGAGGAGAAAGACAGAGCGTTAAGTCCAAATCA 387
Qy 181 AAACCATTTGCTACTGATGAGCGCGATGCCATATAAATTGAAAAAGCTGACTTACTA 240
Db 388 AAACCATTTGCTACTGATGAGCGCGATGCCATATAAATTGAAAAAGCTGACTTACTA 447
Qy 241 AAGCTATTCAAGAACCAATTGATCGCTAAACGTCCACAGTAAACGACACTACTTTGAGT 300
Db 448 AAGCTATTCAAGAACCAATTGATCGCTAAACGTCCACAGTAAACGACACTACTTTGAGT 507
Qy 301 ATTGATTTTGCAGCGATGCAACCAATTAAGTTTAAATTTTGAAGTCTACTTTGCTGAC 360
Db 508 ATTGATTTTGCAGCGATGCAACCAATTAAGTTTAAATTTTGAAGTCTACTTTGCTGAC 567
Qy 361 AAAGATGTTTGGTAAACCTTGCAGCACCACCTGTCGAAGAAATTTTGAAGCGGACAT 420
Db 568 AAAGATGTTTGGTAAACCTTGCAGCACCACCTGTCGAAGAAATTTTGAAGCGGACAT 627
Qy 421 GTGCGGTTAGACCATATAAAGAAAACCAATAAATAAAGCGAAATCTGTTGATGTG 480
Db 628 GTGCGGTTAGACCATATAAAGAAAACCAATAAATAAAGCGAAATCTGTTGATGTG 687
Qy 481 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCGAGTCTCAAA 540
Db 688 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCGAGTCTCAAA 747
Qy 541 GATCTAAGCTATTGAAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 600
Db 748 GATCTAAGCTATTGAAAACCACTAGCTATCGGTGACACCATCACATCTCAAGAAATTA 807
Qy 601 GCTCAAGCAACAAGCAATTTTAAACAAACCCACCCAGGCTATACGATTTATGAACGTGAC 660
Db 808 GCTCAAGCAACAAGCAATTTTAAACAAACCCACCCAGGCTATACGATTTATGAACGTGAC 867
Qy 661 TCCTCAATCTGTCACATCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 720
Db 868 TCCTCAATCTGTCACATCATGACAAATGACATTTTCCGTACGATTTTACCAATGATCAAGAG 927
Qy 721 TTTACTTACCATTGTCGGAACCAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
Db 928 TTTACTTACCATTGTCGGAACCAAGCTTATGAGATCAATAAAAAATCTGGTCTG 987
Qy 781 AATGAGAAATAAACAACACCTGATCTCTGAGAAATATTAAGTCTTAAAAAGGG 840
Db 988 AATGAGAAATAAACAACACCTGATCTCTGAGAAATATTAAGTCTTAAAAAGGG 1047
Qy 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAATACGTT 900

Db 1048 GAAAGCCGATGATCCCTTTGATCGCGAGTCACTTGAACCTGTTACCATCAATACGTT 1107
Qy 901 GATGTCACACCAACGAATGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAAGTAAC 960
Db 1108 GATGTCACACCAACGAATGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAAGTAAC 1167
Qy 961 TTAGACTTCAGAGATTTATACGATCCCTGCTGATGAAGGCTAAACTCTCTACAACTCTC 1020
Db 1168 TTAGACTTCAGAGATTTATACGATCCCTGCTGATGAAGGCTAAACTCTCTACAACTCTC 1227
Qy 1021 GATGCTTTGCTGATTTATGAGTACTACCTTAACCTGGAAGATAGAGATATACGATGAC 1080
Db 1228 GATGCTTTGCTGATTTATGAGTACTACCTTAACCTGGAAGATAGAGATATACGATGAC 1287
Qy 1081 ACCAACCGTATCAATACCGTTTATATGGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
Db 1288 ACCAACCGTATCAATACCGTTTATATGGGCAAGCGACCCGAGGAGATGCTAGCTAT 1347
Qy 1141 CATTTAGCTTATGATTAAGATCGTTTATACCGAAGAGAAACGAGAGTTTACAGCTACCTG 1200
Db 1348 CATTTAGCTTATGATTAAGATCGTTTATACCGAAGAGAAACGAGAGTTTACAGCTACCTG 1407
Qy 1201 CGTTATACAGGACACCTATACCTGATTAACCTGATTAACCTGATTAACCTGATTAACCTG 1245
Db 1408 CGTTATACAGGACACCTATACCTGATTAACCTGATTAACCTGATTAACCTGATTAACCTG 1452

RESULT 10
A20016
LOCUS A20016 1512 bp DNA linear PAT 14-JUL-1995
DEFINITION SEQ ID NO: 24; Nucleotide sequence for streptokinase fused to yeast alpha-factor.
ACCESSION A20016
VERSION A20016.1 GI:1247850
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 1512)
AUTHORS
TITLE PROTEINS AND NUCLEIC ACIDS
JOURNAL Patent: WO 9109125-A 24 27-JUN-1991;
FEATURES Location/Qualifiers
1..1512
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
7..1506
/codon_start=1
/transl_table=11
/product="streptokinase fused to yeast alpha-factor"
/protein_id="CAA01487.1"
/db_xref="GI:1247851"
/translation="MRPFSIFAVLPAASSALAAAPVNTTETETAQIPAEVIGYIDL
EGDFVAVLPFSNNGNLFIINTIASAEKGVSLDKRIAGBEWLLDRPVSNNQ
LVVSAIVTEGNTQDLSLEFIDLTSPAHGKTEQGLSPKSPFADSGAMPKLE
KADLLAQEIQIANVHNDYFEVIDFASDATITDRNGKVFADKGSVTLPTQPVQ
EFLLSGHVRVPRVKEPIQONAKSVIVERTVQPLNDDPFRGLKDTKLKTLAIG
DTITSQELLAAQASILNKTHTPGVTIYVERDSSIVTHNDIIFRTILPMDQSEFTYVKNRE
QAVEINKSGLEENINTDLISEKYVYLKSGKPYDPDRSHLKFETIKYDVNTNEL
LKSQQLTASERNLPDLYDPDKAKGLYNLDAPGIMDYTLTKGVNDHDDTNKRII
TYVMGRKPEGNASVHLAYDKORYTEEREVSYLYRYTGTPIPDNPNDK"

ORIGIN
Query Match 99.7%; Score 1241.8; DB 6; Length 1512;
Best Local Similarity 99.8%; Pred. No. 3.2e-296;
Matches 1243; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGCTCAATCTGTCACAAACAGCAATAGTTGTT 60
Db 262 ATTGCTGGACCTGAGTGGCTGCTAGACCGCTCAATCTGTCACAAACAGCAATAGTTGTT 321

Db 1399 AGCGTTGCTGGTACTGTGTAGGGGAGCAATCAAGACATTAATGTTTAAATTTTGGAAAT 1458
QY 121 GACCTAACATCAGCACTGCTCATGAGGAGAAAGACAGACGCAAGGCTTAAAGTCCAAATCA 180
Db 1459 GACCTAACATCAGCACTGCTCATGAGGAGAAAGACAGACGCAAGGCTTAAAGTCCAAATCA 1518
QY 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAGAGCTGACTTACTA 240
Db 1519 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAGAGCTGACTTACTA 1578
QY 241 AAGGCTATTCAAGAACAAATGATGCTTAAGTCCACAGTAACGACGACTTAAAGTGGTGC 300
Db 1579 AAGGCTATTCAAGAACAAATGATGCTTAAGTCCACAGTAACGACGACTTAAAGTGGTGC 1638
QY 301 ATTGATTTTGAAGCAATGCAACATTAATGATGCGAAACGGCAAGGCTCTACTTTGCTGAC 360
Db 1639 ATTGATTTTGAAGCAATGCAACATTAATGATGCGAAACGGCAAGGCTCTACTTTGCTGAC 1698
QY 361 AAGATGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTGTAAGCGGACAT 420
Db 1699 AAGATGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTGTAAGCGGACAT 1758
QY 421 GTGCGGTTAGACATATAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480
Db 1759 GTGCGGTTAGACATATAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 1818
QY 481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGANTTTCAGACGAGTCTCAAA 540
Db 1819 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGANTTTCAGACGAGTCTCAAA 1878
QY 541 GATACTAAGCTATTGAACACATAGCTATCGGTGACACCAATCAATCTCAAGAAATTA 600
Db 1879 GATACTAAGCTATTGAACACATAGCTATCGGTGACACCAATCAATCTCAAGAAATTA 1938
QY 601 GCTCAAGCAAGCAATTTTAAACAAACCCAGGCTATACGATTTTATGACGCTGAC 660
Db 1939 GCTCAAGCAAGCAATTTTAAACAAACCCAGGCTATACGATTTTATGACGCTGAC 1998
QY 661 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
Db 1999 TCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 2058
QY 721 TTCTACTCATGTCAAAATCGGAAACCAAGCTTATGAGATCAATAAAATCTGCTG 780
Db 2059 TTCTACTCATGTCAAAATCGGAAACCAAGCTTATGAGATCAATAAAATCTGCTG 2118
QY 781 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAGAGG 840
Db 2119 AATGAAGAAATAAACCACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAGAGG 2178
QY 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAATGTTTCAACCAATACGTT 900
Db 2179 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAAATGTTTCAACCAATACGTT 2238
QY 901 GATGTCACCAACCAAGATGCTTAAAGCGAGCAGCTTAAACAGTAGAGCAAGCTAAC 960
Db 2239 GATGTCACCAACCAAGATGCTTAAAGCGAGCAGCTTAAACAGTAGAGCAAGCTAAC 2298
QY 961 TTAGACTTCAGAGATTTATACGATCTCTGATGTAAGCTTAAACTTCTCAACAAATCTC 1020
Db 2299 TTAGACTTCAGAGATTTATACGATCTCTGATGTAAGCTTAAACTTCTCAACAAATCTC 2358
QY 1021 GATGCTTTGGTATTATGGAATACTTAACTGGAAGAGTAGAGGATATCAACGATGAC 1080
Db 2359 GATGCTTTGGTATTATGGAATACTTAACTGGAAGAGTAGAGGATATCAACGATGAC 2418
QY 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGGACCCGAGAGAGATGCTAGCTAT 1140
Db 2419 ACCAACCGTATCATACCGTTTATATGGGCAAGGACCCGAGAGAGATGCTAGCTAT 2478
QY 1141 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAGACGAGAAGTTTACAGCTACCTG 1200

Db 2479 CATTTAGCCTATGATAAAGATCGTTATACCGAAGAAACGAGAGTTTACAGCTACCTG 2538
QY 1201 CGTTATACAGGACACCTATACCTGATTAACCTTAACGACAAATAA 1245
Db 2539 CGTTATACAGGACACCTATACCTGATTAACCTTAACGACAAATAA 2583
RESULT 13
A20030
LOCUS 1467 bp DNA linear PAT 14-JUL-1995
DEFINITION SEQ ID NO: 38; Synthetic nucleotide sequence for streptokinase-hirudin fusion protein.
ACCESSION A20030
VERSION A20030.1 GI:1247865
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1467)
AUTHORS PROTEINS AND NUCLEIC ACIDS
TITLE Patent: WO 9109125-A 38 27-JUN-1991;
JOURNAL Location/Qualifiers
FEATURES
source
1..1467
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 99.5%; Score 1238.8; DB 6; Length 1467;
Best Local Similarity 99.8%; Pred. No. 1.8e-295;
Matches 1240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCAAATTTAGTTGT 60
Db 1 ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACAAACAGCCAAATTTAGTTGT 60
QY 61 AGCGTTGCTGGTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAAAT 120
Db 61 AGCGTTGCTGGTACTGTTGAGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAAAT 120
QY 121 GACCTAACATCAGCACTGCTCATGAGGAGAAAGACAGACGCAAGGCTTAAAGTCCAAATCA 180
Db 121 GACCTAACATCAGCACTGCTCATGAGGAGAAAGACAGACGCAAGGCTTAAAGTCCAAATCA 180
QY 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAGAGCTGACTTACTA 240
Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAAGAGCTGACTTACTA 240
QY 241 AAGGCTATTCAAGAACAAATGATGCTTAACGTCACAGTAAACGACGACTTAAAGTGGTGC 300
Db 241 AAGGCTATTCAAGAACAAATGATGCTTAACGTCACAGTAAACGACGACTTAAAGTGGTGC 300
QY 301 ATTGATTTTGAAGCAATGCAACATTAATGATGCGAAACGGCAAGGCTCTACTTTGCTGAC 360
Db 301 ATTGATTTTGAAGCAATGCAACATTAATGATGCGAAACGGCAAGGCTCTACTTTGCTGAC 360
QY 361 AAGATGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTGTAAGCGGACAT 420
Db 361 AAGATGGTTCGGTAACTTTGCGGACCCAACTGTCCAAGAAATTTTGTGTAAGCGGACAT 420
QY 421 GTGCGGTTAGACATATAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480
Db 421 GTGCGGTTAGACATATAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480
QY 481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGANTTTCAGACGAGTCTCAAA 540
Db 481 GAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGANTTTCAGACGAGTCTCAAA 540
QY 541 GATACTAAGCTATTGAACACATAGCTATCGGTGACACCAATCAATCTCAAGAAATTA 600
Db 541 GATACTAAGCTATTGAACACATAGCTATCGGTGACACCAATCAATCTCAAGAAATTA 600

```
QY 601 GCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGTGAC 660
Db 601 GCTCAAGCACAAAGCATTTTAAACAAAACCCATCCAGGCTATACGATTTATGAACGTGAC 660
QY 661 TCCTCAATCGTCACATCATGACAAATGACAAATTTTCGTAGCATTTTACCAATGGATCAAGAG 720
Db 661 TCCTCAATCGTCACATCATGACAAATGACAAATTTTCGTAGCATTTTACCAATGGATCAAGAG 720
QY 721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780
Db 721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780
QY 781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGG 840
Db 781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGG 840
QY 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACATTTTATGATTTTATGATTTTATGATTTT 900
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACATTTTATGATTTTATGATTTTATGATTTT 900
QY 901 GATGTCACCAACCAAGAAATGCTTAAAGCGAGCAGCTCTTAAACAGCTAGCGAACTAAC 960
Db 901 GATGTCACCAACCAAGAAATGCTTAAAGCGAGCAGCTCTTAAACAGCTAGCGAACTAAC 960
QY 961 TTAGACTTCAGAGATTTATACGATCCTCGTGAAGCTAACTACTCTACAACTCTC 1020
Db 961 TTAGACTTCAGAGATTTATACGATCCTCGTGAAGCTAACTACTCTACAACTCTC 1020
QY 1021 GATGCTTTTGGTATTATGAGACTATACCTTAACTGGAAGTAGAGGATAATCAGATGAC 1080
Db 1021 GATGCTTTTGGTATTATGAGACTATACCTTAACTGGAAGTAGAGGATAATCAGATGAC 1080
QY 1081 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
Db 1081 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
QY 1141 CATTTAGCCTATGATAAGATCGTTTATACCGAAGAACGAGAGATTTACAGTACCTG 1200
Db 1141 CATTTAGCCTATGATAAGATCGTTTATACCGAAGAACGAGAGATTTACAGTACCTG 1200
QY 1201 CGTTATACAGGACACCTATACCTGTATACCTGATACCTCAACGACAAA 1242
Db 1201 CGTTATACAGGACACCTATACCTGTATACCTGATACCTCAACGACAAA 1242

RESULT 14
LOCUS I13218
DEFINITION Sequence 46 from patent US 5434073.
ACCESSION I13218
VERSION I13218.1 GI:910566
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1467)
AUTHORS Dawson,K., Hunter,M.G. and Czaplewski,L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 46 18-JUL-1995;
FEATURES
source 1..1467
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 99.5%; Score 1238.8; DB 6; Length 1467;
Best Local Similarity 99.8%; Pred. No. 1.8e-295;
Matches 1240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATTTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCTCAACACGACCAATTAGTTGTT 60
Db 1 ATTTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCTCAACACGACCAATTAGTTGTT 60
```

```
QY 61 AGCGTTGCTGCTACTGTTGAGGGGACCAATCAAGACACTTAGTCTTAAATTTTTTGAATTT 120
Db 61 AGCGTTGCTGCTACTGTTGAGGGGACCAATCAAGACACTTAGTCTTAAATTTTTTGAATTT 120
QY 121 GACCTAAACATCAGACACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 180
Db 121 GACCTAAACATCAGACACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 180
QY 181 AAACCATTTGCTACTGATAGTGGCGGATGCGCATGCAATTAACCTTGAAGAGCTGACTTACTA 240
Db 181 AAACCATTTGCTACTGATAGTGGCGGATGCGCATGCAATTAACCTTGAAGAGCTGACTTACTA 240
QY 241 AAGGCTATTCAAGAACAAATTTGATCGCTAACAGTCCACAGTAACGACGACTACTTTGAGGTC 300
Db 241 AAGGCTATTCAAGAACAAATTTGATCGCTAACAGTCCACAGTAACGACGACTACTTTGAGGTC 300
QY 301 ATTGATTTTGCAGCGAGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 360
Db 301 ATTGATTTTGCAGCGAGATGCAACCATTTACTGATCGAAACCGCAAGGCTCTACTTTGCTGAC 360
QY 361 AAAGATGTTTCGGTAACTTGGCGACCCACCTGTCGAAGAAATTTTGTAAAGCGGACAT 420
Db 361 AAAGATGTTTCGGTAACTTGGCGACCCACCTGTCGAAGAAATTTTGTAAAGCGGACAT 420
QY 421 GTGCGGTTAGACCATATAAAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
Db 421 GTGCGGTTAGACCATATAAAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
QY 481 GAATATCTCTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAGGCTCTCAAA 540
Db 481 GAATATCTCTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTACAGCAGGCTCTCAAA 540
QY 541 GATACTAAGCTATTGAAACACATGATCTATCGGTGACACCATCATCTCAAGAAATTAATA 600
Db 541 GATACTAAGCTATTGAAACACATGATCTATCGGTGACACCATCATCTCAAGAAATTAATA 600
QY 601 GCTCAAGCAAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGCTGAC 660
Db 601 GCTCAAGCAAAAGCATTTTAAACAAAACCCACCCAGGCTATACGATTTATGAACGCTGAC 660
QY 661 TCCTCAATCGTCACATCATGACAAATTTTCGTAGCATTTTACCAATGGATCAAGAG 720
Db 661 TCCTCAATCGTCACATCATGACAAATTTTCGTAGCATTTTACCAATGGATCAAGAG 720
QY 721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780
Db 721 TTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAATCTGGTCTG 780
QY 781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGG 840
Db 781 AATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTTAAAAAAGG 840
QY 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACATTTTATGATTTTATGATTTTATGATTTT 900
Db 841 GAAAGCCGTATGATCCCTTTGATCGCAGTCACATTTTATGATTTTATGATTTTATGATTTT 900
QY 901 GATGTCACCAACCAAGAAATGCTTAAAGCGAGCAGCTCTTAAACAGCTAGCGAACTAAC 960
Db 901 GATGTCACCAACCAAGAAATGCTTAAAGCGAGCAGCTCTTAAACAGCTAGCGAACTAAC 960
QY 961 TTAGACTTCAGAGATTTATACGATCCTCGTGAAGCTAACTACTCTACAACTCTC 1020
Db 961 TTAGACTTCAGAGATTTATACGATCCTCGTGAAGCTAACTACTCTACAACTCTC 1020
QY 1021 GATGCTTTTGGTATTATGAGACTATACCTTAACTGGAAGTAGAGGATAATCAGATGAC 1080
Db 1021 GATGCTTTTGGTATTATGAGACTATACCTTAACTGGAAGTAGAGGATAATCAGATGAC 1080
QY 1081 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
Db 1081 ACCAACCGTATCAATACCGTTTATATGGCAAGCGACCCGAGGAGAGATGCTAGCTAT 1140
QY 1141 CATTTAGCCTATGATAAGATCGTTTATACCGAAGAACGAGAGATTTACAGTACCTG 1200
```

```
Db 1141 CATTAGCCTATGATAAGATCGTTATACCGAAGAACGAGAAGTTTACAGCTACCTG 1200
Qy 1201 CGTTATACAGGACACCTTATACCTGATAACCCCTACGACAAA 1242
Db 1201 CGTTATACAGGACACCTTATACCTGATAACCCCTACGACAAA 1242

RESULT 15
A20021
LOCUS A20021 2588 bp DNA linear PAT 14-JUL-1995
DEFINITION Seq ID NO: 29; Synthetic nucleotide sequence for
OmpA-streptokinase-streptokinase fusion gene.
ACCESSION A20021
VERSION A20021.1 GI:1247856
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 2588)
AUTHORS
TITLE PROTEINS AND NUCLEIC ACIDS
JOURNAL Patent: WO 9109125-A 29 27-JUN-1991;
FEATURES Location/Qualifiers
source 1..2588
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 99.5%; Score 1238.8; DB 6; Length 2588;
Best Local Similarity 99.8%; Pred. No. 1.8e-295;
Matches 1240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATTGCTGGACCTGAGTGGCTGTAGACGCTCATCTGTCAACACACCAATAGTTGTT 60
Db 67 ATTGCTGGACCTGAGTGGCTGTAGACGCTCATCTGTCAACACACCAATAGTTGTT 126
Qy 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAAT 120
Db 127 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTGGAAAT 186
Qy 121 GACCTAACATCAAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 180
Db 187 GACCTAACATCAAGACCTGCTCATGAGGAGAAAGACAGAGCAAGGCTTAAAGTCCAAATCA 246
Qy 181 AAACATTTGCTACTGATAGTGGCGCGATGCCATAACTTGAAGAAAGCTGACTACTA 240
Db 247 AAACATTTGCTACTGATAGTGGCGCGATGCCATAACTTGAAGAAAGCTGACTACTA 306
Qy 241 AAGGCTATTCAAGAACCAATTGATCGCTTAACGTCCACAGTAACGACGACTACTTTGAGGTC 300
Db 307 AAGGCTATTCAAGAACCAATTGATCGCTTAACGTCCACAGTAACGACGACTACTTTGAGGTC 366
Qy 301 ATTGATTTTGCAGACGATGCAACATTTACTGATGCGAAACGGCAAGGCTCTACTTTGCTGAC 360
Db 367 ATTGATTTTGCAGACGATGCAACATTTACTGATGCGAAACGGCAAGGCTCTACTTTGCTGAC 426
Qy 361 AAAGATGGTTCCGTTAACTTCCGACACCACTGTGCCAGAAATTTTGTCTAAGCGGCAT 420
Db 427 AAAGATGGTTCCGTTAACTTCCGACACCACTGTGCCAGAAATTTTGTCTAAGCGGCAT 486
Qy 421 GTGCGGCTTAGACCATATAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
Db 487 GTGCGGCTTAGACCATATAAGAAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 546
Qy 481 GAATATACTGTACGTTTACTCCCTTAAACCGCTGATGACGATTTTCAGACAGGCTCTCAA 540
Db 547 GAATATACTGTACGTTTACTCCCTTAAACCGCTGATGACGATTTTCAGACAGGCTCTCAA 606
Qy 541 GATTAATAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 600
Db 607 GATTAATAGCTATTGAAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTA 666
```

```
Qy 601 GCTCAAGCACAAGCAATTTTAAACAAACCCACCGCTATACGATTTATGAACGTGAC 660
Db 667 GCTCAAGCACAAGCAATTTTAAACAAACCCATCCAGCTATACGATTTATGAACGTGAC 726
Qy 661 TCCTCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 720
Db 727 TCCTCAATCGTCACTCATGACAATGACATTTTCGTACGATTTTACCAATGGATCAAGAG 786
Qy 721 TTTACTTTACCATGTCAAAATTCGGAAACAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
Db 787 TTTACTTTACCATGTCAAAATTCGGAAACAGCTTATGAGATCAATAAAAAATCTGGTCTG 846
Qy 781 AATGAAGAAATATAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 840
Db 847 AATGAAGAAATATAACACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGGG 906
Qy 841 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTTGAAAACCTGTTCCACATCAATACGTT 900
Db 907 GAAAAGCCGTATGATCCCTTTTGATCGCAGTCACTTTGAAAACCTGTTCCACATCAATACGTT 966
Qy 901 GATGTCAAACCAACCAAGAAATGCTAAAGAGGAGCAGCTCTTAACAGCTAGCGAACGTAAC 960
Db 967 GATGTCAAACCAACCAAGAAATGCTAAAGAGGAGCAGCTCTTAACAGCTAGCGAACGTAAC 1026
Qy 961 TTAGACTTCAGAGATTTTATACGATCCCTCGTGAAGGCTAAACCTACTCTACAAATCTC 1020
Db 1027 TTAGACTTCAGAGATTTTATACGATCCCTCGTGAAGGCTAAACCTACTCTACAAATCTC 1086
Qy 1021 GATGCTTTTGGTATTTATGGAATATACCTTAACTGGAAAAAGTAGAGGATAATCACGATGAC 1080
Db 1087 GATGCTTTTGGTATTTATGGAATATACCTTAACTGGAAAAAGTAGAGGATAATCACGATGAC 1146
Qy 1081 ACCAACCGTATCATACCGGTTTATATGGGCAAGCGCCCGAAGAGAGAGATGCTAGCTAT 1140
Db 1147 ACCAACCGTATCATACCGGTTTATATGGGCAAGCGCCCGAAGAGAGAGATGCTAGCTAT 1206
Qy 1141 CATTTAGCCTTATGATAAAGATCGTTATACCGAAGAACGAGAAGTTTACAGCTACCTG 1200
Db 1207 CATTTAGCCTTATGATAAAGATCGTTATACCGAAGAACGAGAAGTTTACAGCTACCTG 1266
Qy 1201 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAAA 1242
Db 1267 CGTTATACAGGACACCTTATACCTGATTAACCTTAACGACAAA 1308
```

Search completed: May 6, 2004, 06:20:21
Job time : 5093 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 6, 2004, 03:54:48 ; Search time 4627 Seconds
(without alignments)
9467.629 Million cell updates/sec

Title: US-09-882-509-3

Perfect score: 1245

Sequence: 1 attgtggacctgagtggct.....ataacctaacgacaaataa 1245

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main.*

- 1: /cgn2_6/ptodata/2/pna/US06 COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06 COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07 COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US081 COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US082 COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US083 COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US084 COMB.seq.*
- 9: /cgn2_6/ptodata/2/pna/US085 COMB.seq.*
- 10: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
- 11: /cgn2_6/ptodata/2/pna/US087 COMB.seq.*
- 12: /cgn2_6/ptodata/2/pna/US088 COMB.seq.*
- 13: /cgn2_6/ptodata/2/pna/US089 COMB.seq.*
- 14: /cgn2_6/ptodata/2/pna/US090 COMB.seq.*
- 15: /cgn2_6/ptodata/2/pna/US091 COMB.seq.*
- 16: /cgn2_6/ptodata/2/pna/US092A COMB.seq.*
- 17: /cgn2_6/ptodata/2/pna/US092B COMB.seq.*
- 18: /cgn2_6/ptodata/2/pna/US093A COMB.seq.*
- 19: /cgn2_6/ptodata/2/pna/US093B COMB.seq.*
- 20: /cgn2_6/ptodata/2/pna/US094 COMB.seq.*
- 21: /cgn2_6/ptodata/2/pna/US095A COMB.seq.*
- 22: /cgn2_6/ptodata/2/pna/US095B COMB.seq.*
- 23: /cgn2_6/ptodata/2/pna/US095C COMB.seq.*
- 24: /cgn2_6/ptodata/2/pna/US095D COMB.seq.*
- 25: /cgn2_6/ptodata/2/pna/US096A COMB.seq.*
- 26: /cgn2_6/ptodata/2/pna/US096B COMB.seq.*
- 27: /cgn2_6/ptodata/2/pna/US096C COMB.seq.*
- 28: /cgn2_6/ptodata/2/pna/US096D COMB.seq.*
- 29: /cgn2_6/ptodata/2/pna/US096E COMB.seq.*
- 30: /cgn2_6/ptodata/2/pna/US097A COMB.seq.*
- 31: /cgn2_6/ptodata/2/pna/US097B COMB.seq.*
- 32: /cgn2_6/ptodata/2/pna/US097C COMB.seq.*
- 33: /cgn2_6/ptodata/2/pna/US098A COMB.seq.*
- 34: /cgn2_6/ptodata/2/pna/US098B COMB.seq.*
- 35: /cgn2_6/ptodata/2/pna/US098C COMB.seq.*
- 36: /cgn2_6/ptodata/2/pna/US098D COMB.seq.*
- 37: /cgn2_6/ptodata/2/pna/US099A COMB.seq.*
- 38: /cgn2_6/ptodata/2/pna/US099B COMB.seq.*
- 39: /cgn2_6/ptodata/2/pna/US099C COMB.seq.*
- 40: /cgn2_6/ptodata/2/pna/US099D COMB.seq.*
- 41: /cgn2_6/ptodata/2/pna/US099E COMB.seq.*
- 42: /cgn2_6/ptodata/2/pna/US099F COMB.seq.*
- 43: /cgn2_6/ptodata/2/pna/US100A COMB.seq.*

- 44: /cgn2_6/ptodata/2/pna/US100B COMB.seq.*
- 45: /cgn2_6/ptodata/2/pna/US101A COMB.seq.*
- 46: /cgn2_6/ptodata/2/pna/US101B COMB.seq.*
- 47: /cgn2_6/ptodata/2/pna/US102A COMB.seq.*
- 48: /cgn2_6/ptodata/2/pna/US102B COMB.seq.*
- 49: /cgn2_6/ptodata/2/pna/US103A COMB.seq.*
- 50: /cgn2_6/ptodata/2/pna/US103B COMB.seq.*
- 51: /cgn2_6/ptodata/2/pna/US104A COMB.seq.*
- 52: /cgn2_6/ptodata/2/pna/US104B COMB.seq.*
- 53: /cgn2_6/ptodata/2/pna/US106 COMB.seq.*
- 54: /cgn2_6/ptodata/2/pna/US107A COMB.seq.*
- 55: /cgn2_6/ptodata/2/pna/US107B COMB.seq.*
- 56: /cgn2_6/ptodata/2/pna/US6000 COMB.seq.*
- 57: /cgn2_6/ptodata/2/pna/US6001 COMB.seq.*
- 58: /cgn2_6/ptodata/2/pna/US6002 COMB.seq.*
- 59: /cgn2_6/ptodata/2/pna/US6003 COMB.seq.*
- 60: /cgn2_6/ptodata/2/pna/US6004 COMB.seq.*
- 61: /cgn2_6/ptodata/2/pna/US6005 COMB.seq.*
- 62: /cgn2_6/ptodata/2/pna/US6006 COMB.seq.*
- 63: /cgn2_6/ptodata/2/pna/US6007 COMB.seq.*
- 64: /cgn2_6/ptodata/2/pna/US6008 COMB.seq.*
- 65: /cgn2_6/ptodata/2/pna/US6009 COMB.seq.*
- 66: /cgn2_6/ptodata/2/pna/US6010 COMB.seq.*
- 67: /cgn2_6/ptodata/2/pna/US6011 COMB.seq.*
- 68: /cgn2_6/ptodata/2/pna/US6012 COMB.seq.*
- 69: /cgn2_6/ptodata/2/pna/US6013 COMB.seq.*
- 70: /cgn2_6/ptodata/2/pna/US6014 COMB.seq.*
- 71: /cgn2_6/ptodata/2/pna/US6015 COMB.seq.*
- 72: /cgn2_6/ptodata/2/pna/US6016 COMB.seq.*
- 73: /cgn2_6/ptodata/2/pna/US6017 COMB.seq.*
- 74: /cgn2_6/ptodata/2/pna/US6018 COMB.seq.*
- 75: /cgn2_6/ptodata/2/pna/US6019 COMB.seq.*
- 76: /cgn2_6/ptodata/2/pna/US6020 COMB.seq.*
- 77: /cgn2_6/ptodata/2/pna/US6021 COMB.seq.*
- 78: /cgn2_6/ptodata/2/pna/US6022 COMB.seq.*
- 79: /cgn2_6/ptodata/2/pna/US6023A COMB.seq.*
- 80: /cgn2_6/ptodata/2/pna/US6023B COMB.seq.*
- 81: /cgn2_6/ptodata/2/pna/US6024 COMB.seq.*
- 82: /cgn2_6/ptodata/2/pna/US6025 COMB.seq.*
- 83: /cgn2_6/ptodata/2/pna/US6026 COMB.seq.*
- 84: /cgn2_6/ptodata/2/pna/US6027 COMB.seq.*
- 85: /cgn2_6/ptodata/2/pna/US6028 COMB.seq.*
- 86: /cgn2_6/ptodata/2/pna/US6029 COMB.seq.*
- 87: /cgn2_6/ptodata/2/pna/US6030 COMB.seq.*
- 88: /cgn2_6/ptodata/2/pna/US6031 COMB.seq.*
- 89: /cgn2_6/ptodata/2/pna/US6032 COMB.seq.*
- 90: /cgn2_6/ptodata/2/pna/US6033 COMB.seq.*
- 91: /cgn2_6/ptodata/2/pna/US6034 COMB.seq.*
- 92: /cgn2_6/ptodata/2/pna/US6035 COMB.seq.*
- 93: /cgn2_6/ptodata/2/pna/US6036 COMB.seq.*
- 94: /cgn2_6/ptodata/2/pna/US6037 COMB.seq.*
- 95: /cgn2_6/ptodata/2/pna/US6038 COMB.seq.*
- 96: /cgn2_6/ptodata/2/pna/US6039 COMB.seq.*
- 97: /cgn2_6/ptodata/2/pna/US6040 COMB.seq.*
- 98: /cgn2_6/ptodata/2/pna/US6041 COMB.seq.*
- 99: /cgn2_6/ptodata/2/pna/US6042 COMB.seq.*
- 100: /cgn2_6/ptodata/2/pna/US6043 COMB.seq.*
- 101: /cgn2_6/ptodata/2/pna/US6044 COMB.seq.*
- 102: /cgn2_6/ptodata/2/pna/US6045 COMB.seq.*
- 103: /cgn2_6/ptodata/2/pna/US6046 COMB.seq.*
- 104: /cgn2_6/ptodata/2/pna/US6047 COMB.seq.*
- 105: /cgn2_6/ptodata/2/pna/US6048 COMB.seq.*
- 106: /cgn2_6/ptodata/2/pna/US6049 COMB.seq.*
- 107: /cgn2_6/ptodata/2/pna/US6050 COMB.seq.*
- 108: /cgn2_6/ptodata/2/pna/US6051 COMB.seq.*
- 109: /cgn2_6/ptodata/2/pna/US6052 COMB.seq.*
- 110: /cgn2_6/ptodata/2/pna/US6053 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Db 481 GAATATAGTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGTCTCAAA 540
 QY 541 GATACCTAGCTATTGAACCACTAGCTATCGGTGACACCATCCTCAAGATTTACTTA 600
 Db 541 GATACCTAGCTATTGAACCACTAGCTATCGGTGACACCATCCTCAAGATTTACTTA 600
 QY 601 GCTCAAGCAAAAGCAATTTTAAACCAAAACCCCAAGGCTATACGATTTTATGAACGTGAC 660
 Db 601 GCTCAAGCAAAAGCAATTTTAAACCAAAACCCCAAGGCTATACGATTTTATGAACGTGAC 660
 QY 661 TCCTCAATCGTCACTCATGACATGACATTTTCCTGATCGATTTTACCAATGATCAAGAG 720
 Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCCTGATCGATTTTACCAATGATCAAGAG 720
 QY 721 TTTTACTTACGATGTCAAAATCGGAAACAAAGCTTATGAGATCAATATAAAAAATCTGCTCTG 780
 Db 721 TTTTACTTACGATGTCAAAATCGGAAACAAAGCTTATGAGATCAATATAAAAAATCTGCTCTG 780
 QY 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAAGGG 840
 Db 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAAGGG 840
 QY 841 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAACATCAAAATACGTT 900
 Db 841 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAACATCAAAATACGTT 900
 QY 901 GATGTCACCAACCAAGAAATGCTTAAAGCGAGCAGCTCTTAAACAGTACGGAACGTAAC 960
 Db 901 GATGTCACCAACCAAGAAATGCTTAAAGCGAGCAGCTCTTAAACAGTACGGAACGTAAC 960
 QY 961 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGGCTTAAACTACTCTCAACAATCTC 1020
 Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGGCTTAAACTACTCTCAACAATCTC 1020
 QY 1021 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAGATGAGAGTAATCAGGATGAC 1080
 Db 1021 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAGATGAGAGTAATCAGGATGAC 1080
 QY 1081 ACCAACCGTATCATACCCGTTTATATGGGCAAGCAGACCCGAGGAGAGATGCTAGCTAT 1140
 Db 1081 ACCAACCGTATCATACCCGTTTATATGGGCAAGCAGACCCGAGGAGAGATGCTAGCTAT 1140
 QY 1141 CATTTAGCTTATGATTAAGATCGTTTATACCGAAGAAAGACGAGAGTTTACGACTACCTG 1200
 Db 1141 CATTTAGCTTATGATTAAGATCGTTTATACCGAAGAAAGACGAGAGTTTACGACTACCTG 1200
 QY 1201 CGTTTATACGAGGACCTATACCTGATACCCCTTAAACGACAAATAA 1245
 Db 1201 CGTTTATACGAGGACCTATACCTGATACCCCTTAAACGACAAATAA 1245

RESULT 2

US-09-882-509-3
 ; Sequence 3, Application US/09882509
 ; GENERAL INFORMATION:
 ; APPLICANT: Kuppasamy, Mosuvan
 ; APPLICANT: Srinivas, Vellimedu K
 ; APPLICANT: Lahiri, Subhra
 ; APPLICANT: Ella, Krishna
 ; APPLICANT: Khatri, Ghan S
 ; TITLE OF INVENTION: Recombinant Streptokinase
 ; FILE REFERENCE: 51321.003
 ; CURRENT APPLICATION NUMBER: US/09/882,509
 ; CURRENT FILING DATE: 2001-06-15
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 1245
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis (ATCC 9542)
 US-09-882-509-3

Query Match 100.0%; Score 1245; DB 36; Length 1245;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATTTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACCAAGCCAAATTTAGTTGT 60
 Db 1 ATTTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACCAAGCCAAATTTAGTTGT 60
 QY 61 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTTTGAATTT 120
 Db 61 AGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTTAGTCTTAAATTTTTTGAATTT 120
 QY 121 GACCTAAACATCAAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 180
 Db 121 GACCTAAACATCAAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 180
 QY 181 AAACCATTTGCTGATGATGAGTGGCGGATGCCACATAAATCTTGAAGAGCTGACTTACTA 240
 Db 181 AAACCATTTGCTGATGATGAGTGGCGGATGCCACATAAATCTTGAAGAGCTGACTTACTA 240
 QY 241 AAGGCTATTCAAGAAACAAATTTGATCGTCAACGTCACAGTAACGACGACTTCTTGAAGT 300
 Db 241 AAGGCTATTCAAGAAACAAATTTGATCGTCAACGTCACAGTAACGACGACTTCTTGAAGT 300
 QY 301 ATTGATTTGCAAGCGATGCAACCAATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 360
 Db 301 ATTGATTTGCAAGCGATGCAACCAATTTACTGATCGAAACGGCAAGGCTTACTTTGCTGAC 360
 QY 361 AAAGATGTTTGGTAAACCTTGGCCGACCCCAACCTGTCGAAGAAATTTTGTAAAGCGACAT 420
 Db 361 AAAGATGTTTGGTAAACCTTGGCCGACCCCAACCTGTCGAAGAAATTTTGTAAAGCGACAT 420
 QY 421 GTGCGGTTAGACCATATAAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480
 Db 421 GTGCGGTTAGACCATATAAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480
 QY 481 GAATATCTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 540
 Db 481 GAATATCTGTACGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACGAGTCTCAA 540
 QY 541 GATACTAAGCTATTGAACCACTAGCTATCGGTGACCATCACAATCTCAAGAAATTTACTA 600
 Db 541 GATACTAAGCTATTGAACCACTAGCTATCGGTGACCATCACAATCTCAAGAAATTTACTA 600
 QY 601 GCTCAAGCAAAAGCAATTTTAAACCAAAACCCCAAGGCTATACGATTTTATGAACGTGAC 660
 Db 601 GCTCAAGCAAAAGCAATTTTAAACCAAAACCCCAAGGCTATACGATTTTATGAACGTGAC 660
 QY 661 TCCTCAATCGTCACTCATGACATGACATTTTCCTGATCGATTTTACCAATGATCAAGAG 720
 Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCCTGATCGATTTTACCAATGATCAAGAG 720
 QY 721 TTTTACTTACCATGTCAAAATCGGAAACAAAGCTTATGAGATCAATATAAAAAATCTGCTCTG 780
 Db 721 TTTTACTTACCATGTCAAAATCGGAAACAAAGCTTATGAGATCAATATAAAAAATCTGCTCTG 780
 QY 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAAGGG 840
 Db 781 AATGAAGAAATAAACCAACTGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAAGGG 840
 QY 841 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAACATCAAAATACGTT 900
 Db 841 GAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTCAACATCAAAATACGTT 900
 QY 901 GATGTCACCAACCAAGAAATGCTTAAAGCGAGCAGCTCTTAAACAGTACGGAACGTAAC 960
 Db 901 GATGTCACCAACCAAGAAATGCTTAAAGCGAGCAGCTCTTAAACAGTACGGAACGTAAC 960
 QY 961 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGGCTTAAACTACTCTCAACAATCTC 1020
 Db 961 TTAGACTTCAGAGATTTATACGATCTCTGATTAAGGCTTAAACTACTCTCAACAATCTC 1020
 QY 1021 GATGCTTTTGGTATTATGACATATACCTTAACTGGAAGATGAGAGTAATCAGGATGAC 1080

Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGTAGAGGATAATCACGATGAC 1080
 Qy 1081 ACCAACGGTATCATACCGTTTATATGGGCAAGCGACCCGAGAGAGATGCTAGCTAT 1140
 Db 1081 ACCAACGGTATCATACCGTTTATATGGGCAAGCGACCCGAGAGAGATGCTAGCTAT 1140
 Qy 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAAAGAGAGATTTACAGCTACCTG 1200
 Db 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAAAGAGAGATTTACAGCTACCTG 1200
 Qy 1201 CGTTATACAGGACACCTTATACCTGATAACCCCTAAACGACAAATAA 1245
 Db 1201 CGTTATACAGGACACCTTATACCTGATAACCCCTAAACGACAAATAA 1245

RESULT 3

US-09-471-349-1
 ; Sequence 1, Application US/09471349
 ; GENERAL INFORMATION:
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Roy, Chaiti
 ; APPLICANT: Rajagopal, Kammara
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE PROTEINS POSSESSING ALTERED
 ; TITLE OF INVENTION: PLASMINOGEN ACTIVATION CHARACTERISTICS AND A PROCESS FOR THE
 ; FILE REFERENCE: 07064/009001
 ; CURRENT APPLICATION NUMBER: US/09/471.349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1245
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1242)
 US-09-471-349-1

Query Match 98.3%; Score 1224.2; DB 20; Length 1245;
 Best Local Similarity 99.0%; Pred. No. 0; Mismatches 13; Indels 0; Gaps 0;
 Matches 1232; Conservative 0;
 Qy 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTAGTTGTT 60
 Db 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCCAATTAGTTGTT 60
 Qy 61 AGCGTTGCTGCTACTGTTGAGGAGCAAGATCAAGACATTAGTCTTAATTTTGAATTT 120
 Db 61 AGCGTTGCTGCTACTGTTGAGGAGCAAGATCAAGACATTAGTCTTAATTTTGAATTT 120
 Qy 121 GACCTAACATCACCACCTGCTCATGAGGAAAGACAGAGCAAGCTTAAGTCCAAATCA 180
 Db 121 GATCTAACATCAGCACTGCTCATGAGGAAAGACAGAGCAAGCTTAAGTCCAAATCA 180
 Qy 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAACTTGAATAAGCTGACTTACTA 240
 Db 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAACTTGAATAAGCTGACTTACTA 240
 Qy 241 AAGGCTATTCAAGACATTTGCTTAAGTCCACAGTCCACAGTACGCTACTTTGAGTTC 300
 Db 241 AAGGCTATTCAAGACATTTGCTTAAGTCCACAGTCCACAGTACGCTACTTTGAGTTC 300
 Qy 301 ATTGATTTTGAAGCGATGCAACCACTTACTGATGAAACGCAAGGTCTACTTTGCTGAC 360
 Db 301 ATTGATTTTGAAGCGATGCAACCACTTACTGATGAAACGCAAGGTCTACTTTGCTGAC 360

Qy 361 AAAGATGGTTCGGTAACCTTGGCGACCCCAACCTGTCCAAAGAAATTTTGTCTAAGCGACAT 420
 Db 361 AAAGATGGTTCGGTAACCTTGGCGACCCCAACCTGTCCAAAGAAATTTTGTCTAAGCGACAT 420
 Qy 421 GTGCGGTTAGACCATATATAAGAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
 Db 421 GTGCGGTTAGACCATATATAAGAAACCAATACAAAATCAAGCGAAATCTGTTGATGTG 480
 Qy 481 GAATATACGTGTACAGTTTACTCCCTTAAACCCCTGTATGACGATTTTCAGACAGGTCTCAAA 540
 Db 481 GAATATACGTGTACAGTTTACTCCCTTAAACCCCTGTATGACGATTTTCAGACAGGTCTCAAA 540
 Qy 541 GATACCTAAGCTATTGAAAAACACTAGCTATCCGTTAAGCGATTTTACGATTTTACGATTTTAC 600
 Db 541 GATACCTAAGCTATTGAAAAACACTAGCTATCCGTTAAGCGATTTTACGATTTTACGATTTTAC 600
 Qy 601 GCTCAAGCACAAGCAATTTTAAACCAACCCAGGCTATACGATTTTATGAACTGAC 660
 Db 601 GCTCAAGCACAAGCAATTTTAAACCAACCCAGGCTATACGATTTTATGAACTGAC 660
 Qy 661 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 720
 Db 661 TCCTCAATCGTCACTCATGACATGACATTTTCGTTACGATTTTACCAATGGATCAAGAG 720
 Qy 721 TTTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCAATATAAATACTGCTGCTG 780
 Db 721 TTTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCAATATAAATACTGCTGCTG 780
 Qy 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGCTTCTTAAAAAGGG 840
 Db 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACGCTTCTTAAAAAGGG 840
 Qy 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCCACCTCAATATACGTT 900
 Db 841 GAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTCCACCTCAATATACGTT 900
 Qy 901 GATGTCACACCAACGAATGCTTAAAGAGGAGCAGCTCTTAAACGATAGCGAACGTAAC 960
 Db 901 GATGTCACACCAACGAATGCTTAAAGAGGAGCAGCTCTTAAACGATAGCGAACGTAAC 960
 Qy 961 TTAGACTTTCAGAGATTTTATACGATCTCTGTAAGCTTAACTCTTCAACAATCTC 1020
 Db 961 TTAGACTTTCAGAGATTTTATACGATCTCTGTAAGCTTAACTCTTCAACAATCTC 1020
 Qy 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGTAGAGGATATCAATCTC 1080
 Db 1021 GATGCTTTTGGTATTATGACTATACCTTAACTGGAAGTAGAGGATATCAATCTC 1080
 Qy 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGAGAGATGCTAGCTAT 1140
 Db 1081 ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGAGAGATGCTAGCTAT 1140
 Qy 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAAAGAGAGATTTACAGTACTCTG 1200
 Db 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAAAGAGAGATTTACAGTACTCTG 1200
 Qy 1201 CGTTATACAGGACACCTTATACCTGATAACCCCTAAACGACAAATAA 1245
 Db 1201 CGTTATACAGGACACCTTATACCTGATAACCCCTAAACGACAAATAA 1245

RESULT 4

US-09-940-235-1
 ; Sequence 1, Application US/09940235
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chaiti
 ; APPLICANT: Rajagopal, Kammara
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE

; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/09/940,235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1245
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1242)
 US-09-940-235-1

Query Match 98.3%; Score 1224.2; DB 39; Length 1245;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 1232; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY	1	ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTAGTTGTT	60
DB	1	ATTGCTGGACCTCAGTGGCTGTAGACCGTCCATCTGTCAACACAGCCCAATTTAGTTGTT	60
QY	61	AGCGTTGCTGGTACTGTGTGAGGGGAGCGAATCAAGACATTTAGTTTAAATTTTGAAT	120
DB	61	AGCGTTGCTGGTACTGTGTGAGGGGAGCGAATCAAGACATTTAGTTTAAATTTTGAAT	120
QY	121	GACCTAACATCAGACCTGTCTGTGAGGAGAAACACAGAGCGTTTAACTTCAAAATCA	180
DB	121	GATCTAACATCAGACCTGTCTGTGAGGAGAAACACAGAGCGTTTAACTTCAAAATCA	180
QY	181	AAACCAATTTGCTACTGTAGTGGCGGATGCCATATAAATCTGAAAAGCTGACTTACTA	240
DB	181	AAACCAATTTGCTACTGTAGTGGCGGATGCCATATAAATCTGAAAAGCTGACTTACTA	240
QY	241	AAGGCTATTCAAGAACAAATTTGATCGTTAAGTCCACAGTAAACGACTACTTTGAGGTC	300
DB	241	AAGGCTATTCAAGAACAAATTTGATCGTTAAGTCCACAGTAAACGACTACTTTGAGGTC	300
QY	301	ATTGATTTTGAAGCGATGCAACATTTACTGATCGAAGCGAAGTCTACTTTGCTGAC	360
DB	301	ATTGATTTTGAAGCGATGCAACATTTACTGATCGAAGCGAAGTCTACTTTGCTGAC	360
QY	361	AAAGATGGTTTGGTAACTTGGCGACCCCAACCTGTCCAGAAATTTTGTAAAGCGGACAT	420
DB	361	AAAGATGGTTTGGTAACTTGGCGACCCCAACCTGTCCAGAAATTTTGTAAAGCGGACAT	420
QY	421	GTGGCGTTTGAACCATATAAAGAAAACCAATACAAATCAAGGAAATCTGTTGATGTG	480
DB	421	GTGGCGTTTGAACCATATAAAGAAAACCAATACAAATCAAGGAAATCTGTTGATGTG	480
QY	481	GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGGTTCTCAA	540
DB	481	GAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTTTCAAGGTTCTCAA	540
QY	541	GATATTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA	600
DB	541	GATATTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCATCTCAAGAAATTTACTA	600
QY	601	GCTCAAGCACAAAGCAATTTTAAACCAAAACCCACCGCTATAGATTTTAAAGCTGAC	660
DB	601	GCTCAAGCACAAAGCAATTTTAAACCAAAACCCACCGCTATAGATTTTAAAGCTGAC	660
QY	661	TCCTCAATCGTCACTGACAAATGACATTTTCCGTAGATTTTACCAATGGATCAAGAG	720
DB	661	TCCTCAATCGTCACTGACAAATGACATTTTCCGTAGATTTTACCAATGGATCAAGAG	720

QY	721	TTTACTTACCATGTCAAAATCGGAAACAGCTTATGAGATCAATAAAAAATCTGCTCG	780
DB	721	TTTACTTACCATGTCAAAATCGGAAACAGCTTATGAGATCAATAAAAAATCTGCTCG	780
QY	781	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACCTTTAAAAAGGG	840
DB	781	AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACCTTTAAAAAGGG	840
QY	841	GAAGAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAGTGTTCACCATCAATACGTT	900
DB	841	GAAGAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAGTGTTCACCATCAATACGTT	900
QY	901	GATGTCAACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAACGTAAC	960
DB	901	GATGTCAACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAACGTAAC	960
QY	961	TTAGACTTCAGAGATTTATACGATCCCTGATGATAGGCTTAACTTCTTACCAATCTC	1020
DB	961	TTAGACTTCAGAGATTTATACGATCCCTGATGATAGGCTTAACTTCTTACCAATCTC	1020
QY	1021	GATGCTTTGTTATGATGACTATACCTTAACTGGGAAAAGTAGAGGATTAATCAGGATGAC	1080
DB	1021	GATGCTTTGTTATGATGACTATACCTTAACTGGGAAAAGTAGAGGATTAATCAGGATGAC	1080
QY	1081	ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGAGAGAGAAATGCTAGCTAT	1140
DB	1081	ACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGAGAGAGAAATGCTAGCTAT	1140
QY	1141	CATTAGCCTATGATTAAGATCGTTATACGAAAGAGACGAGAACTTTTACAGCTACCTG	1200
DB	1141	CATTAGCCTATGATTAAGATCGTTATACGAAAGAGACGAGAACTTTTACAGCTACCTG	1200
QY	1201	CGTTATACAGGACACTATACCTGATTAACCTTAAACCGCTTAAACGACAAATAA	1245
DB	1201	CGTTATACAGGACACTATACCTGATTAACCTTAAACCGCTTAAACGACAAATAA	1245

RESULT 5

US-10-631-558-1

Sequence 1, Application US/10631558

GENERAL INFORMATION:

APPLICANT: Kumar, Rajesh

APPLICANT: Sahni, Girish

APPLICANT: Roy, Chait

APPLICANT: Rajagopal, Kammar

APPLICANT: Nihalani, Deepak

APPLICANT: Sundaram, Vasudha

APPLICANT: Yadav, Mahavir

TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE

TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION OF SAID

TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID

TITLE OF INVENTION: PROTEIN

FILE REFERENCE: 07064-009002

CURRENT APPLICATION NUMBER: US/10/631,558

PRIOR FILING DATE: 2003-07-31

PRIOR APPLICATION NUMBER: US/09/940,235

PRIOR FILING DATE: 2002-04-09

PRIOR APPLICATION NUMBER: 09/471,349

PRIOR FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: IN 3825/DEL/98

PRIOR FILING DATE: 1998-12-24

NUMBER OF SEQ ID NOS: 28

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 1245

TYPE: DNA

ORGANISM: Streptococcus equisimilis

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(1242)

US-10-631-558-1

Query Match

98.3%; Score 1224.2; DB 53; Length 1245;

Best Local Similarity 99.0%; Pred. No. 0;
Matches 1232; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATTAAGTTGTT 60
Db |||
Qy 1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATTAAGTTGTT 60
Db |||
Qy 61 AGCGTTGCTGCTACTGTTTGGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAAT 120
Db |||
Qy 61 AGCGTTGCTGCTACTGTTTGGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAAT 120
Db |||
Qy 121 GACCTAACATCAGACCTGCTCTGTGAGGAGAAAGACAGAGCGCTTAAGTCCAAATCA 180
Db |||
Qy 121 GATCTAACATCAGACCTGCTCTGTGAGGAGAAAGACAGAGCGCTTAAGTCCAAATCA 180
Db |||
Qy 181 AAACCATTTGCTACTGATAGTGGCGGATGCCATTAATCTGAAAGCTGACTTACTA 240
Db |||
Qy 181 AAACCATTTGCTACTGATAGTGGCGGATGCCATTAATCTGAAAGCTGACTTACTA 240
Db |||
Qy 241 AAGGCTATTCAAGAACAAATTGATCGCTAACCTGTCACAGTAAACGAGCTACTTTGAGTC 300
Db |||
Qy 241 AAGGCTATTCAAGAACAAATTGATCGCTAACCTGTCACAGTAAACGAGCTACTTTGAGTC 300
Db |||
Qy 301 ATTGATTTGCAAGGATGCAACATTAATCTGATGAAACGCAAGGCTACTTTGCTGAC 360
Db |||
Qy 301 ATTGATTTGCAAGGATGCAACATTAATCTGATGAAACGCAAGGCTACTTTGCTGAC 360
Db |||
Qy 361 AAAGATGTTGCGTAACTGCGGACCACTGTCACAGTAAATTTTGTAAAGCGGACAT 420
Db |||
Qy 361 AAAGATGTTGCGTAACTGCGGACCACTGTCACAGTAAATTTTGTAAAGCGGACAT 420
Db |||
Qy 421 GTGCGGTTAGACCATATAAGAAACCAATAACAATCAAGCGAAATCTGTTGATGTG 480
Db |||
Qy 421 GTGCGGTTAGACCATATAAGAAACCAATAACAATCAAGCGAAATCTGTTGATGTG 480
Db |||
Qy 481 GAATATCTGACGTTTACTCCCTTAAACCTGATGACGATTCAGACCAAGGCTCTCAA 540
Db |||
Qy 481 GAATATCTGACGTTTACTCCCTTAAACCTGATGACGATTCAGACCAAGGCTCTCAA 540
Db |||
Qy 541 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTA 600
Db |||
Qy 541 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACTCTCAAGAAATTA 600
Db |||
Qy 601 GCTCAAGCAACAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACGTCAC 660
Db |||
Qy 601 GCTCAAGCAACAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAACGTCAC 660
Db |||
Qy 661 TCCTCAATCGTCACTATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAG 720
Db |||
Qy 661 TCCTCAATCGTCACTATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAG 720
Db |||
Qy 721 TTTTACTTACCATGTCAAAATTCGGGAAACAAGCTTATGAGATCAATTAATAAATCTGCTG 780
Db |||
Qy 721 TTTTACTTACCATGTCAAAATTCGGGAAACAAGCTTATGAGATCAATTAATAAATCTGCTG 780
Db |||
Qy 781 AATGAAGAAATAACAAACACTGATCTGTGAGAAATATTAAGTCTTAAATAAGG 840
Db |||
Qy 781 AATGAAGAAATAACAAACACTGATCTGTGAGAAATATTAAGTCTTAAATAAGG 840
Db |||
Qy 841 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 900
Db |||
Qy 841 GAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTCAACCATCAATACGTT 900
Db |||
Qy 901 GATGTCAACACCAAGCAATTTGTAAGGAGCAAGCTCTTAAACAGCTAGGCAAGCTTAAC 960
Db |||
Qy 901 GATGTCAACACCAAGCAATTTGTAAGGAGCAAGCTCTTAAACAGCTAGGCAAGCTTAAC 960
Db |||
Qy 961 TTAGACTTCAGAGATTTTATACGATCCTCGTGAAGGCTTAACTACTCTACAAATCTC 1020
Db |||
Qy 961 TTAGACTTCAGAGATTTTATACGATCCTCGTGAAGGCTTAACTACTCTACAAATCTC 1020
Db |||
Qy 1021 GATGCTTTTGGTATTATGCACTATACCTTAACTGGAAGTAGAGGATTAACGATGAC 1080
Db |||

Db 1021 GATGCTTTTGGTATTATGGACTATACCTTAACGTGAAAGTAGAGGATAATCAGATGAC 1080
Qy 1081 ACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGAGAGAAATGCTAGCTAT 1140
Db |||
Qy 1081 ACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAAGAGAGAAATGCTAGCTAT 1140
Db |||
Qy 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAAAGAGAGTTTACAGCTACCTG 1200
Db |||
Qy 1141 CATTTAGCTTATGATAAAGATCGTTTATACCGAAGAGAAAGAGAGTTTACAGCTACCTG 1200
Db |||
Qy 1201 CGTTATACAGGACACCTTATACCTGATAACCCCTAAACGCAAAATAA 1245
Db |||
Qy 1201 CGTTATACAGGACACCTTATACCTGATAACCCCTAAACGCAAAATAA 1245
Db |||

RESULT 6
US-09-471-349-5
; Sequence 5, Application US/09471349
; GENERAL INFORMATION:
; APPLICANT: Sahni, Girish
; APPLICANT: Kumar, Rajesh
; APPLICANT: Roy, Chaiti
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE PROTEINS POSSESSING ALTERED
; TITLE OF INVENTION: PLASMINOGEN ACTIVATION CHARACTERISTICS AND A PROCESS FOR THE
; FILE REFERENCE: 07064/009001
; CURRENT APPLICATION NUMBER: US/09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-471-349-5

Query Match 98.2%; Score 1222.6; DB 20; Length 1377;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATTAAGTTGTT 60
Db |||
Qy 133 ATTGCTGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCAATTAAGTTGTT 192
Db |||
Qy 61 AGCGTTGCTGCTACTGTTTGGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAAT 120
Db |||
Qy 193 AGCGTTGCTGCTACTGTTTGGGGGAGCAATCAAGACATTAAGTCTTAAATTTTGAAT 252
Db |||
Qy 121 GACCTAACATCAGACCTGCTCTGTGAGGAGAAAGACAGAGCGCTTAAAGTCCAAATCA 180
Db |||
Qy 253 GATCTAACATCAGACCTGCTCTGTGAGGAGAAAGACAGAGCGCTTAAAGTCCAAATCA 312
Db |||
Qy 181 AAACCATTTGCTACTGATAGTGGCGGATGCCATTAATCTGAAAGCTGACTTACTA 240
Db |||
Qy 313 AAACCATTTGCTACTGATAGTGGCGGATGCCATTAATCTGAAAGCTGACTTACTA 372
Db |||
Qy 241 AAGGCTATTCAAGAACAAATTGATCGCTTAAACAGTAAACGAGCTACTTTGAGTC 300
Db |||
Qy 373 AAGGCTATTCAAGAACAAATTGATCGCTTAAACAGTAAACGAGCTACTTTGAGTC 432
Db |||
Qy 301 ATTGATTTGCAAGGATGCAACATTAATCTGATGAAACGCAAGGCTACTTTGCTGAC 360
Db |||
Qy 433 ATTGATTTGCAAGGATGCAACATTAATCTGATGAAACGCAAGGCTACTTTGCTGAC 492
Db |||
Qy 361 AAAGATGTTGCGTAACTGCGGACCACTGTCACAGTAAATTTTGTAAAGCGGACAT 420
Db |||
Qy 493 AAAGATGTTGCGTAACTGCGGACCACTGTCACAGTAAATTTTGTAAAGCGGACAT 552
Db |||

Db 913 AATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGGG 972
QY 841 GAAAAGCGGTATGATCCCTTTGATCGAGTCACTTTGAAAGTGTTCACCATCAAAATACGTT 900
Db 973 GAAAAGCGGTATGATCCCTTTGATCGAGTCACTTTGAAAGTGTTCACCATCAAAATACGTT 1032
QY 901 GATGCAACACCAAGAAATGCTTAAAGCGAGAGCTCTTAAACAGCTAGCGAAAGCTAAC 960
Db 1033 GATGCGATACCAAGAAATGCTTAAAGCGAGAGCTCTTAAACAGCTAGCGAAAGCTAAC 1092
QY 961 TTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTTAACCTCTACAGCAATCTC 1020
Db 1093 TTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTTAACCTCTACAGCAATCTC 1152
QY 1021 GATGCTTTTGGTATTTATGAGACTATACCTTAACTGGAAGAGTAGAGATAATCACGATGAC 1080
Db 1153 GATGCTTTTGGTATTTATGAGACTATACCTTAACTGGAAGAGTAGAGATAATCACGATGAC 1212
QY 1081 ACCAACCGTATCATAAACCGTTTATATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1213 ACCAACCGTATCATAAACCGTTTATATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1272
QY 1141 CATTTAGCTTATGATTAAG 1200
Db 1273 CATTTAGCTTATGATTAAG 1332
QY 1201 CGTTATACAGGACACCTATACCTGATTAACCTTAACTGGAAGAGTAGAGATAATCACGATGAC 1245
Db 1333 CGTTATACAGGACACCTATACCTGATTAACCTTAACTGGAAGAGTAGAGATAATCACGATGAC 1377

RESULT 8

US-10-631-558-5
; Sequence 5, Application US/10631558
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sami, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-10-631-558-5

Query Match 98.2%; Score 1222.6; DB 53; Length 1377;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 ATTGCTGACCTGAGTGGTCTAGACCGTCACTCTGCAACAAAGAGAGAGAGAGAGAGAGAGAG 60
Db 133 ATTGCTGACCTGAGTGGTCTAGACCGTCACTCTGCAACAAAGAGAGAGAGAGAGAGAGAGAG 192
QY 61 AGCGTTGCTGAGTGGTCTAGACCGTCACTCTGCAACAAAGAGAGAGAGAGAGAGAGAGAGAG 120

Db 193 AGCGTTGCTGAGTGGTCTAGACCGTCACTCTGCAACAAAGAGAGAGAGAGAGAGAGAGAGAG 252
QY 121 GACCTAACATCAGCAGCTGCTCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 253 GATCTAACATCAGCAGCTGCTCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312
QY 181 AAACCAATTTGCTACTGATAGTGGCGGATGCGACATAAACTTTGAAAGAGCTGACTTACTA 240
Db 313 AAACCAATTTGCTACTGATAGTGGCGGATGCGACATAAACTTTGAAAGAGCTGACTTACTA 372
QY 241 AAGGCTATTCAAGAACCAATTTGATCGCTAAAGCTGCGACAGTAAAGAGAGAGAGAGAGAGAGAG 300
Db 373 AAGGCTATTCAAGAACCAATTTGATCGCTAAAGCTGCGACAGTAAAGAGAGAGAGAGAGAGAGAG 432
QY 301 ATTGATTTTGGCAAGCGATGCAACCAATTTGATCGCTAAAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 433 ATTGATTTTGGCAAGCGATGCAACCAATTTGATCGCTAAAGAGAGAGAGAGAGAGAGAGAGAG 492
QY 361 AAGATGGTTTGGTAAACCTTTGCGGACCGAACCTCTGCAAGAAATTTTGGCTAAGCGGAGAT 420
Db 493 AAGATGGTTTGGTAAACCTTTGCGGACCGAACCTCTGCAAGAAATTTTGGCTAAGCGGAGAT 552
QY 421 GTGCGCGTTAGACCATATTAAGAAACCAATTTGATCGCTAAAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 553 GTGCGCGTTAGACCATATTAAGAAACCAATTTGATCGCTAAAGAGAGAGAGAGAGAGAGAGAGAG 612
QY 481 GAATATATCTGATACAGTTTACTCTCCCTTAAACCTGATGACGATTTTCAAGACAGAGTCTCAAA 540
Db 613 GAATATATCTGATACAGTTTACTCTCCCTTAAACCTGATGACGATTTTCAAGACAGAGTCTCAAA 672
QY 541 GATATAGCTATTGAAACCACTAGCTATCGGTGACACCATCAATCTCAAGAAATTTACTA 600
Db 673 GATATAGCTATTGAAACCACTAGCTATCGGTGACACCATCAATCTCAAGAAATTTACTA 732
QY 601 GCTCAAGCACAAAGCAATTTTAAACCAACCCAGGCTATACGATTTATGAGAGAGAGAGAGAGAGAG 660
Db 733 GCTCAAGCACAAAGCAATTTTAAACCAACCCAGGCTATACGATTTATGAGAGAGAGAGAGAGAGAG 792
QY 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
Db 793 TCCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 852
QY 721 TTTTACTTACATGTCACAAATTCGGGAAACAGCTTATGAGATCAATATAAAATCTGGTCTG 780
Db 853 TTTTACTTACATGTCACAAATTCGGGAAACAGCTTATGAGATCAATATAAAATCTGGTCTG 912
QY 781 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGAGAGAGAGAGAG 840
Db 913 AATGAAGAAATAAACAACACCTGATCTCTGAGAAATATTACGCTCTTAAAAAGAGAGAGAGAGAG 972
QY 841 GAAAAGCGGTATGATCCCTTTGATCGGAGTCACTTTGAAACCTGTTTCAACATCAATACGTT 900
Db 973 GAAAAGCGGTATGATCCCTTTGATCGGAGTCACTTTGAAACCTGTTTCAACATCAATACGTT 1032
QY 901 GATGTCACACCAAGCAATTTGCTTAAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 1033 GATGTCACACCAAGCAATTTGCTTAAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1092
QY 961 TTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTTAACCTCTACAGCAATCTC 1020
Db 1093 TTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTTAACCTCTACAGCAATCTC 1152
QY 1021 GATGCTTTTGGTATTTATGAGACTATACCTTAACTGGAAGAGTAGAGATAATCACGATGAC 1080
Db 1153 GATGCTTTTGGTATTTATGAGACTATACCTTAACTGGAAGAGTAGAGATAATCACGATGAC 1212
QY 1081 ACCAACCGTATCATAAACCGTTTATATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1213 ACCAACCGTATCATAAACCGTTTATATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1272
QY 1141 CATTTAGCTTATGATTAAG 1200
Db 1273 CATTTAGCTTATGATTAAG 1332


```

; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-11

Query Match      98.2%; Score 1222.6; DB 39; Length 1782;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACACGCCAATAGTTGTT 60
Db 538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACACGCCAATAGTTGTT 597
QY 61 AGCGTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAAAT 120
Db 598 AGCGTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAAAT 657
QY 121 GACCTACATCAGACCTGCTCATGGAGGAAAGACAGACGAAGCTTAAGTCCAAATCA 180
Db 658 GATCTACATCAGACCTGCTCATGGAGGAAAGACAGACGAAGCTTAAGTCCAAATCA 717
QY 181 AAACCATTTGCTACTGATAGTGGCGGATGCGACATAAACTTGAAAAAGCTGACTTACTA 240
Db 718 AAACCATTTGCTACTGATAGTGGCGGATGCGACATAAACTTGAAAAAGCTGACTTACTA 777
QY 241 AAGCTATTCAGAAACAAATTGATCGCTAAACGTCACAGTAAGACGACTACTTTGAGTGC 300
Db 778 AAGCTATTCAGAAACAAATTGATCGCTAAACGTCACAGTAAGACGACTACTTTGAGTGC 837
QY 301 ATTGATTTTGAAGCGATGCAACCACTTACTGATCGAAACGCGAAGTCTTACTTTGCTGAC 360
Db 838 ATTGATTTTGAAGCGATGCAACCACTTACTGATCGAAACGCGAAGTCTTACTTTGCTGAC 897
QY 361 AAAGATGTTTGGTAACTTCCGACCCCAACCTGTCACAAAGTATTTTGTAAAGCGGACAT 420
Db 898 AAAGATGTTTGGTAACTTCCGACCCCAACCTGTCACAAAGTATTTTGTAAAGCGGACAT 957
QY 421 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGGAATCTGTTGATGTG 480
Db 958 GTGCGGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGGAATCTGTTGATGTG 1017
QY 481 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCAGGTTCTCAA 540
Db 1018 GAATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTTCAAGCAGGTTCTCAA 1077
QY 541 GATATAGCTTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 600
Db 1078 GATATAGCTTATTGAAACACTAGCTATCGGTGACACCATCATCATCTCAAGAAATTA 1137
QY 601 GCTCAAGCACAAGCATTTTAAACAAACCCACCGCTATAGATTATGACGCTGAC 660
Db 1138 GCTCAAGCACAAGCATTTTAAACAAACCCACCGCTATAGATTATGACGCTGAC 1197
QY 661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTAGCATTTTACCAATCGATCAAGAG 720
Db 1198 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTAGCATTTTACCAATCGATCAAGAG 1257
QY 721 TTTACTTACCATTGTCAAAATTCGGGAAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
Db 1258 TTTACTTACCATTGTCAAAATTCGGGAAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 1317
QY 781 AATGAAGAAATAAACAACACTGACCTGATCTGAGAAATATTAGCTCTTAAAAAAGGG 840
Db 1318 AATGAAGAAATAAACAACACTGACCTGATCTGAGAAATATTAGCTCTTAAAAAAGGG 1377
QY 841 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAATGTTTCAACCATCAATACGTT 900
Db 1378 GAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTTGAAATGTTTCAACCATCAATACGTT 1437

```

RESULT 11

```

US-10-631-558-11
; Sequence 11, Application US/10631558
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahn, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-10-631-558-11

```

```

Query Match      98.2%; Score 1222.6; DB 53; Length 1782;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1231; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACACGCCAATAGTTGTT 60
Db 538 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTCTCAACACACGCCAATAGTTGTT 597
QY 61 AGCGTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAAAT 120
Db 598 AGCGTTCTGGTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGGAAAT 657
QY 121 GACCTACATCAGACCTGCTCATGGAGGAAAGACAGACGAAGCTTAAAGTCCAAATCA 180

```

Db 658 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 717
QY 181 AAACCAATTTGCTACTGATAGTGGGGGATGCGACATAAACTTGAAGAGCTGACTACTA 240
Db 718 AAACCAATTTGCTACTGATAGTGGGGGATGCGACATAAACTTGAAGAGCTGACTACTA 777
QY 241 AAGCGTATTCAAGAAACAATGATCGCTTAACGCTCCACAGTAACGACGACTACTTTGAGGTC 300
Db 778 AAGCGTATTCAAGAAACAATGATCGCTTAACGCTCCACAGTAACGACGACTACTTTGAGGTC 837
QY 301 ATTGATTTTCAAGCGATGCAACCATTAATGATGCGAAACGCGAAGGTCTACTTTGCTGAC 360
Db 838 ATTGATTTTCAAGCGATGCAACCATTAATGATGCGAAACGCGAAGGTCTACTTTGCTGAC 897
QY 361 AAAGATGTTTCGGTAAACCTTTCGCGACCAACCTGTCGAAGAAATTTTGTCTAAGCGACAT 420
Db 898 AAAGATGTTTCGGTAAACCTTTCGCGACCAACCTGTCGAAGAAATTTTGTCTAAGCGACAT 957
QY 421 GTGCGGTTAGACCATATAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480
Db 958 GTGCGGTTAGACCATATAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 1017
QY 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCTCAA 540
Db 1018 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACCAGGTCTCAA 1077
QY 541 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATCTACTA 600
Db 1078 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATCTACTA 1137
QY 601 GCTCAAGCAAAAGCATTTTAAACCAAAACCCAGGCTATACGATTTATGAAAGCTGAC 660
Db 1138 GCTCAAGCAAAAGCATTTTAAACCAAAACCCAGGCTATACGATTTATGAAAGCTGAC 1197
QY 661 TCCTCAATCGTCTCATGACAAATGACATTTTCGTTACGATTTTACCAATGATCAAGAG 720
Db 1198 TCCTCAATCGTCTCATGACAAATGACATTTTCGTTACGATTTTACCAATGATCAAGAG 1257
QY 721 TTTTACTTACCATGTCAAAATCGGAAACAGCTTATGAGATCAATAAAATCTGCTG 780
Db 1258 TTTTACTTACCGTGTAAATTCGGAAACAGCTTATGAGATCAATAAAATCTGCTG 1317
QY 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 840
Db 1318 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAAGG 1377
QY 841 GAAAAGCGGTATGATCCCTTTGATTCGAGTCACTTGAATCTGTTCAACATCAATACGTT 900
Db 1378 GAAAAGCGGTATGATCCCTTTGATTCGAGTCACTTGAATCTGTTCAACATCAATACGTT 1437
QY 901 GATGTCAACACCAAGAAATGCTTAAAGAGGAGCGAGCTCTTAAACAGCTAGCGAAGCTAAC 960
Db 1438 GATGTGATACCAAGAAATGCTTAAAGAGGAGCGAGCTCTTAAACAGCTAGCGAAGCTAAC 1497
QY 961 TTAGACTTCAGAGATTTATACGATCCCTGATGAAAGGCTAAACTACTCTCAACAATCTC 1557
Db 1498 TTAGACTTCAGAGATTTATACGATCCCTGATGAAAGGCTAAACTACTCTCAACAATCTC 1557
QY 1021 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTATGAGGATATACGATGAC 1080
Db 1558 GATGCTTTTGGTATTATGGAATATACCTTAACTGGAAGAGTATGAGGATATACGATGAC 1617
QY 1081 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCGAAGGAGAGATGCTAGCTAT 1140
Db 1618 ACCAACCGTATCATACCGTTTATATGCGCAAGCGACCGAAGGAGAGATGCTAGCTAT 1677
QY 1141 CATTTAGCTTATGATAAAGATCGTTATACCGAAGAGAACGAGAGTTTACAGTACCTG 1200
Db 1678 CATTTAGCTTATGATAAAGATCGTTATACCGAAGAGAACGAGAGTTTACAGTACCTG 1737
QY 1201 CGTTATACAGGACCACTTATACCTGTATACGCTTACGCAAAATTA 1245

Db 1738 CGTTATACAGGACACCTTATACCTGTATACCTTAAAGCAAAATAA 1782
RESULT 12
US-09-569-920-2
; Sequence 2, Application US/09569920
; GENERAL INFORMATION:
; APPLICANT: Reddy, Vemuri B.
; APPLICANT: Lether, Ethan
; TITLE OF INVENTION: VASODILATOR-THROMBOLYTIC FUSION PROTEIN
; TITLE OF INVENTION: AND CONJUGATES
; FILE REFERENCE: 10284-026001
; CURRENT APPLICATION NUMBER: US/09/569,920
; CURRENT FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
US-09-569-920-2
Query Match 98.1%; Score 1221.2; DB 24; Length 1254;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1229; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATAGTTGTT 60
Db 7 ATTGCTGGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACAGCAATAGTTGTT 66
QY 61 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAAT 120
Db 67 AGCGTTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTTAAATTTTGAAT 126
QY 121 GACCTAATCATCAGACCTGCTGATGAGGAAACAGACAGCAAGCTTAAGTCCAAATCA 180
Db 127 GATCTAATCATCAGACCTGCTGATGAGGAAACAGACAGCAAGCTTAAGTCCAAATCA 186
QY 181 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTCAAAAAGCTGACTACTA 240
Db 187 AAACCAATTTGCTACTGATAGTGGCGGATGCCACATAAACTTCAAAAAGCTGACTACTA 246
QY 241 AAGCGTATTCAAGAAACAATGATCGCTTAACGCTCCACAGTAACGACGCTACTTTAGGTC 300
Db 247 AAGCGTATTCAAGAAACAATGATCGCTTAACGCTCCACAGTAACGACGCTACTTTAGGTC 306
QY 301 ATTGATTTTGAAGCAAGTCAACCATTAAGTCAAGAAACGCGAAGGTCTACTTTGCTGAC 360
Db 307 ATTGATTTTGAAGCAAGTCAACCATTAAGTCAAGAAACGCGAAGGTCTACTTTGCTGAC 366
QY 361 AAAGATGTTTCGGTAAACCTTTCGCGACCAACCTGTCGAAGAAATTTTGTCTAAGCGACAT 420
Db 367 AAAGATGTTTCGGTAAACCTTTCGCGACCAACCTGTCGAAGAAATTTTGTCTAAGCGACAT 426
QY 421 GTGCGGTTAGACCATATAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 480
Db 427 GTGCGGTTAGACCATATAAGAAACCAATCAAAATCAAGCGAAATCTGTTGATGTG 486
QY 481 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGAGGATTTTCAGACCAGGTCTCAA 540
Db 487 GAATATCTGTACAGTTTACTCCCTTAAACCCCTGATGAGGATTTTCAGACCAGGTCTCAA 546
QY 541 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATCTACTA 600
Db 547 GATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATCTACTA 606
QY 601 GCTCAAGCAAAAGCATTTTAAACCAAAACCCAGGCTATACGATTTATGAAAGCTGAC 660
Db 607 GCTCAAGCAAAAGCATTTTAAACCAAAACCCAGGCTATACGATTTATGAAAGCTGAC 666
QY 661 TCCTCAATCGTCTCATGACATTAAGTCTGATGAGTCTTACCAATGATCAAGAG 720
Db 667 TCCTCAATCGTCTCATGACATTAAGTCTTACCAATGATCAAGAG 726

```
QY 721 TTTACTTACCATGTCAAAAATCGGAAACAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
Db 727 TTTACTTACCGTGTAAATAATCGGAAACAGCTTATAGGATCAATAAAAAATCTGGTCTG 786
QY 781 AATGAAGAAATAAACAACACATGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAGGG 840
Db 787 AATGAAGAAATAAACAACACATGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAGGG 846
QY 841 GAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAACCATCAATATCGTT 900
Db 847 GAAAGCGGTATGATCCCTTTGATCGGAGTCACTTGAACCTGTTCAACCATCAATATCGTT 906
QY 901 GATGTCAACACCAACGAATCTCTAAAGACGAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 960
Db 907 GATGTCAACACCAACGAATCTCTAAAGAGTGAAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 966
QY 961 TTAGACTTCAGAGATTTATAGATCTCTGATGAGGCTAACTACTCTACAACAATCTC 1020
Db 967 TTAGACTTCAGAGATTTATAGATCTCTGATGAGGCTAACTACTCTACAACAATCTC 1026
QY 1021 GATGCTTTTGGTATATGAGCTATACCTTAACTGGAAGTAGAGGATAAATCAGATGAC 1080
Db 1027 GATGCTTTTGGTATATGAGCTATACCTTAACTGGAAGTAGAGGATAAATCAGATGAC 1086
QY 1081 ACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAATGCTAGCTAT 1140
Db 1087 ACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAATGCTAGCTAT 1146
QY 1141 CATTTAGCTATGATAAGATCGTTTATACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1147 CATTTAGCTATGATAAGATCGTTTATACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1206
QY 1201 CGTTATACAGGACACCTTATACCTGATACCTTAAACCTTAAACGACAAA 1242
Db 1207 CGTTATACAGGACACCTTATACCTGATACCTTAAACCTTAAACGACAAA 1248

RESULT 13
US-09-569-920-3
; Sequence 3, Application US/09569920
; GENERAL INFORMATION:
; APPLICANT: Reddy, Vemuri B.
; APPLICANT: Lerner, Ethan
; TITLE OF INVENTION: VASODILATOR-THROMBOLYTIC FUSION PROTEIN
; TITLE OF INVENTION: AND CONJUGATES
; FILE REFERENCE: 10284-026001
; CURRENT APPLICATION NUMBER: US/09/569,920
; CURRENT FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 8893
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-09-569-920-3

Query Match 98.1%; Score 1221.2; DB 24; Length 8893;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1229; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 ATTTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGAGCAATAGTGTGT 60
Db 5935 ATTTGCTGACCTGAGTGGCTGTAGACCGTCCATCTGTCAACACGAGCAATAGTGTGT 5994
QY 61 AGCGTTGCTGGTACTGTTGAGGGGAGAGATCAAGACATTAAGTCTTAAATTTTGAATTT 120
Db 5995 AGCGTTGCTGGTACTGTTGAGGGGAGAGATCAAGACATTAAGTCTTAAATTTTGAATTT 6054
QY 121 GACCTAACATCAGACCTCTCATGAGGAAGAGACAGCAGGCTTAAGTCCAAATCA 180
```

```
Db 6055 GATCTAACATCAAGACCTGCTCATGAGGAAAGACAGAGCAAGGCTTAAGTCCAAATCA 6114
QY 181 AAACCAATTTCTCTACTGATAGTGGCGGATGCCACATAAACTTGAAAAAGCTGACTTACTA 240
Db 6115 AAACCAATTTCTCTACTGATAGTGGCGGATGCCACATAAACTTGAAAAAGCTGACTTACTA 6174
QY 241 AAGGCTATTCAAGAACAAATTTGATCGCTAACGCTCCACAGTAGTAACGACGACTACTTTGAGGTC 300
Db 6175 AAGGCTATTCAAGAACAAATTTGATCGCTAACGCTCCACAGTAGTAACGACGACTACTTTGAGGTC 6234
QY 301 ATTGATTTTCAAGCGATGCAACCAATTTACTGATGCGAAACGCGAAGGCTTACTTTGCTGAC 360
Db 6235 ATTGATTTTCAAGCGATGCAACCAATTTACTGATGCGAAACGCGAAGGCTTACTTTGCTGAC 6294
QY 361 AAAGATGTTTCGGTAAACCTTTGCGGACCAACCTGTCCTCAAGAAATTTTGTCTAAGCGGACAT 420
Db 6295 AAAGATGTTTCGGTAAACCTTTGCGGACCAACCTGTCCTCAAGAAATTTTGTCTAAGCGGACAT 6354
QY 421 GTGCGGCTTAGACCATATAAGAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
Db 6355 GTGCGGCTTAGACCATATAAGAAACCAATACAAACCAAGCGAAATCTGTTGATGTG 6414
QY 481 GAATATATCTGACAGTTTACTCTCCCTTAAACCTGATGACGATTTTCAGACGAGTCTCAAA 540
Db 6415 GAATATATCTGACAGTTTACTCTCCCTTAAACCTGATGACGATTTTCAGACGAGTCTCAAA 6474
QY 541 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACTACATCTCAAGAAATTTACTA 600
Db 6475 GATACTAAGCTATTGAAAAACACTAGCTATCGGTGACACCACTACATCTCAAGAAATTTACTA 6534
QY 601 GCTCAAGCAACAAAGCATTTTAAACAAACCCAGCGCTATACGTTTATGAACGTCGAC 660
Db 6535 GCTCAAGCAACAAAGCATTTTAAACAAACCCAGCGCTATACGTTTATGAACGTCGAC 6594
QY 661 TCCTCAATCGTCACTCATGACAAATGACATTTTCGTAGCATTTTACCAATGGATCAAGAG 720
Db 6595 TCCTCAATCGTCACTCATGACAAATGACATTTTTCGTAGCATTTTACCAATGGATCAAGAG 6654
QY 721 TTTACTTTACCATGTCAAAAATCGGAAACAAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
Db 6655 TTTACTTTACCGTGTAAAAATCGGAAACAAAGCTTATAGGATCAATAAAAAATCTGGTCTG 6714
QY 781 AATGAAGAAATAAACAACACATGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAGGG 840
Db 6715 AATGAAGAAATAAACAACACATGACCTGATCTCTGAGAAATATTAAGTCTTTAAAAAGGG 6774
QY 841 GAAAAGCGGTATGATCCCTTTGATCGGAGTCACTTTGAAAACCTGTTCAACCATCAATATAGCTT 900
Db 6775 GAAAAGCGGTATGATCCCTTTGATCGGAGTCACTTTGAAAACCTGTTCAACCATCAATATAGCTT 6834
QY 901 GATGTCAACACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 960
Db 6835 GATGTCAACACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAAACGTAAC 6894
QY 961 TTAGACTTCAGAGATTTATAGCTCCCTGATGAGGCTAACTACTCTACAAATCTC 1020
Db 6895 TTAGACTTCAGAGATTTATAGCTCCCTGATGAGGCTAACTACTCTACAAATCTC 6954
QY 1021 GATGCTTTTGGTATTTATGGACTATACCTTAACTGGAAGTAGAGGATAAATCAGATGAC 1080
Db 6955 GATGCTTTTGGTATTTATGGACTATACCTTAACTGGAAGTAGAGGATAAATCAGATGAC 7014
QY 1081 ACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAATGCTAGCTAT 1140
Db 7015 ACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGGAAGGAGAGAATGCTAGCTAT 7074
QY 1141 CATTTAGCTATGATAAGATCGTTTATACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 7075 CATTTAGCTATGATAAGATCGTTTATACGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7134
QY 1201 CGTTATACAGGACACCTTATACCTTAAACCTTAAACCTTAAACGACAAA 1242
Db 7135 CGTTATACAGGACACCTTATACCTTAAACCTTAAACGACAAA 7176
```

RESULT 14
US-09-471-349-6
; Sequence 6, Application US/09471349
; GENERAL INFORMATION:
; APPLICANT: Sahni, Girish
; APPLICANT: Kumar, Rajesh
; APPLICANT: Roy, Chaiti
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE PROTEINS POSSESSING ALTERED
; TITLE OF INVENTION: PLASMINOGEN ACTIVATION CHARACTERISTICS AND A PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF SAID PROTEINS
; FILE REFERENCE: 07064/009001
; CURRENT APPLICATION NUMBER: US/09/471,349
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated primer
US-09-471-349-6

Query Match 97.2%; Score 1209.8; DB 20; Length 1327;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1223; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 1 ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCTCAACAGCCCAATTAGTTGTT 60
DB 83 ATAGCTGGTCTCAATGGCTACTAGATCGTCTTCTGTAATAACAGCCCAATTGGTTGTT 142
QY 61 AGCTTGTCTGGTACTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAAT 120
DB 143 AGCTTGTCTGGTACTGTGAGGGGAGCAATCAAGACATTAGTCTTAAATTTTTTGAATC 202
QY 121 GACCTAACATCAGCACTGCTCATGAGGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 180
DB 203 GATCTAACATCAGCACTGCTCATGAGGAAAGACAGAGCAAGCTTAAAGTCCAAATCA 262
QY 181 AAACCAATTTGCTACTGATAGTGGGGGATGCCACATAAACTTGAAAAGCTGACTTACTA 240
DB 263 AAACCAATTTGCTACTGATAGTGGGGGATGCCACATAAACTTGAAAAGCTGACTTACTA 322
QY 241 AAGCTATTCAAGACAAATTTGATCGCTTAAGCTCCACAGTAAAGCACTACTTTGAGGTC 300
DB 323 AAGCTATTCAAGACAAATTTGATCGCTTAAGCTCCACAGTAAAGCACTACTTTGAGGTC 382
QY 301 ATTGATTTTGAAGCGATGCAACCATTTACTGATGCGAAAGCAAGGCTTACTTTGCTGAC 360
DB 383 ATTGATTTTGAAGCGATGCAACCATTTACTGATGCGAAAGCAAGGCTTACTTTGCTGAC 442
QY 361 AAGATGGTTGGTAACTTTGCGGACCAACCTGTCACAGAAATTTTGTCAAGCGACAT 420
DB 443 AAGATGGTTGGTAACTTTGCGGACCAACCTGTCACAGAAATTTTGTCAAGCGGACAT 502
QY 421 GTGGCGGTAGACCATATAAGAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
DB 503 GTGGCGGTAGACCATATAAGAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 562
QY 481 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 540
DB 563 GAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTCAGACAGGCTCTCAA 622
QY 541 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCTCAAGAAATTTACTA 600

DB 623 GATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAAATTTACTA 682
QY 601 GCTCAAGCACAAGCATTTTAAACAAAACCCAGGCTATACGATTTTATGAACCTGAC 660
DB 683 GCTCAAGCACAAGCATTTTAAACAAAACCCAGGCTATACGATTTTATGAACCTGAC 742
QY 661 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 720
DB 743 TCCTCAATCGTCACTCATGACCAATGACATTTTCCGTACGATTTTACCAATGGATCAAGAG 802
QY 721 TTTACTTACCATGTCAAAATCGGGAACAGCTTATGAGATCAATAAAAATCTGCTCTG 780
DB 803 TTTACTTACCGTGTAAAAAATCGGGAACAGCTTATGAGATCAATAAAAAATCTGCTCTG 862
QY 781 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACCTCTTAAAAAAGG 840
DB 863 AATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTTACCTCTTAAAAAAGG 922
QY 841 GAAAAGCCGTATGATCCCTTTTGTATCGCAGTCACTTTGAAACTGTTTCAACATCAATACGTT 900
DB 923 GAAAAGCCGTATGATCCCTTTTGTATCGCAGTCACTTTGAAACTGTTTCAACATCAATACGTT 982
QY 901 GATGTCACACCAACGAAATTTCTAAAAAGCGAGCAGCTCTTAAACAGCTAGCGAACCTAAC 960
DB 983 GATGTCGATACCAACGAAATTTCTAAAAAGTGAGCAGCTCTTAAACAGCTAGCGAACCTAAC 1042
QY 961 TTAGACTTCAGAGATTTTATACGATCTCTGATGAAAGCTAAACTTCTTACAACTATCTC 1020
DB 1043 TTAGACTTCAGAGATTTTATACGATCTCTGATGAAAGCTAAACTTCTTACAACTATCTC 1102
QY 1021 GATGCTTTTGGTATTTATGACATATACCTTAACTGGAAGAGTAGAGGATATATCAGATGAC 1080
DB 1103 GATGCTTTTGGTATTTATGACATATACCTTAACTGGAAGAGTAGAGGATATATCAGATGAC 1162
QY 1081 ACCAACCGTATCATACCGTCTTATATGGCAAGCGACCGAGAGAGAGATGCTAGCTAT 1140
DB 1163 ACCAACCGTATCATACCGTCTTATATGGCAAGCGACCGAGAGAGAGATGCTAGCTAT 1222
QY 1141 CATTTAGCCCTATGATAAAGATCGTTATATCGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1223 CATTTAGCCCTATGATAAAGATCGTTATATCGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1282
QY 1201 CGTTATACAGGACACCTATACCTGATAACCTTAAACGACAAATAA 1245
DB 1283 CGTTATACAGGACACCTATACCTGATAACCTTAAACGACAAATAA 1327

RESULT 15

US-09-940-235-6
; Sequence 6, Application US/09940235
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chaiti
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327

```
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-6

Query Match      97.2%; Score 1209.8; DB 39; Length 1327;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1223; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1 ATTCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACACAGCCAAATTAGTTGTT 60
Db 83 ATAGCTGGTCTGAATGGCTACTAGATCGTCTTCTGTAAATAACAGCCAAATTGTTGTT 142

Qy 61 AGCGTGTGCTGCTAGTGTGGAGGACGAAATCAAGACATTAAGTCTTAATTTTGAATTT 120
Db 143 AGCGTGTGCTGCTAGTGTGGAGGACGAAATCAAGACATTAAGTCTTAATTTTGAATTT 202

Qy 121 GACCTAACATCAGCACTGCTCATGAGGAAAGACAGAGCAAGCTTAAGTCCAAAATCA 180
Db 203 GATCTAACATCAGCACTGCTCATGAGGAAAGACAGAGCAAGCTTAAGTCCAAAATCA 262

Qy 181 AAACCATTTGCTACTGATPAGTGGCGGATGCCACATAACTTTGAAAAGCTGACTTACTA 240
Db 263 AAACCATTTGCTACTGATPAGTGGCGGATGCCACATAACTTTGAAAAGCTGACTTACTA 322

Qy 241 AAGGCTATTCAAGAACAAATGATCGCTAACTGCTCAAGTCAAGTCAAGGACTACTTTGAGGTC 300
Db 323 AAGGCTATTCAAGAACAAATGATCGCTAACTGCTCAAGTCAAGTCAAGGACTACTTTGAGGTC 382

Qy 301 ATTGATTTTGAAGGATGCAACCATTAATCTGATGAAACGGCAAGTCTACTTTGCTGAC 360
Db 383 ATTGATTTTGAAGGATGCAACCATTAATCTGATGAAACGGCAAGTCTACTTTGCTGAC 442

Qy 361 AAAGATGTTTCGGTAACTTTCGCCGACCCCAACCTGTCCAAAGAAATTTTGCTAAGCGGACAT 420
Db 443 AAAGATGTTTCGGTAACTTTCGCCGACCCCAACCTGTCCAAAGAAATTTTGCTAAGCGGACAT 502

Qy 421 GTGGCGGTAGACCATATAAAGAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 480
Db 503 GTGGCGGTAGACCATATAAAGAAAACCAATACAAATCAAGCGAAATCTGTTGATGTG 562

Qy 481 GAATATCTGTACGATTTTACTTCCCTTAAACCTGATGACGATTTTACAGACAGGCTCAAA 540
Db 563 GAATATCTGTACGATTTTACTTCCCTTAAACCTGATGACGATTTTACAGACAGGCTCAAA 622

Qy 541 GATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAATA 600
Db 623 GATCTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCAATCTCAAGAAATTAATA 682

Qy 601 GCTCAAGCACAAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAAAGTGTGAC 660
Db 683 GCTCAAGCACAAAGCAATTTTAAACAAACCCAGGCTATACGATTTATGAAAGTGTGAC 742

Qy 661 TCCTCAATCGTCACATCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 720
Db 743 TCCTCAATCGTCACATCATGACAAATGACATTTTCGGTACGATTTTACCAATGGATCAAGAG 802

Qy 721 TTTTACTTACCAGTCAAAAATCGGGAAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 780
Db 803 TTTTACTTACCAGTCAAAAATCGGGAAACAAGCTTATGAGATCAATAAAAAATCTGGTCTG 862

Qy 781 AATGAAGAAATAAACAACATGACCTGATCTCTGAGAAATATTAGCTCTTAAABAAGGG 840
Db 863 AATGAAGAAATAAACAACATGACCTGATCTCTGAGAAATATTAGCTCTTAAABAAGGG 922

Qy 841 GAAAAGCGGTATGATCCCTTTTGTGCGCAGTCACTTTGAAACTGTTCAACCATCAAAATACGTT 900
Db 923 GAAAAGCGGTATGATCCCTTTTGTGCGCAGTCACTTTGAAACTGTTCAACCATCAAAATACGTT 982

Qy 901 GATGTCAACCAACGAATTTGCTAAAAACGAGGACAGCTCTTTAAACAGCTAGCGAAAGTAAAC 960
Db 983 GATGTCAACCAACGAATTTGCTAAAAACGAGGACAGCTCTTTAAACAGCTAGCGAAAGTAAAC 1042

Qy 961 TTAGACTTCAGAGATTTTATACGATCCTCGTGATTAAGGCTAAACTACTCTACAAATCTC 1020
```

Search completed: May 6, 2004, 07:44:45
Job time : 4633 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 6, 2004, 03:28:18 ; Search time 290 Seconds
(without alignments)
3661.179 Million cell updates/sec

Title: US-09-882-509-3

Perfect score: 1245

Sequence: 1 attgtcgaccctgagtgctt.....ataaccctaagacaataa 1245

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1493552 segs, 426402401 residues

Total number of hits satisfying chosen parameters: 2987104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA New.*

- 1: /cgn2_6/prodata/1/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/prodata/1/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/prodata/1/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/prodata/1/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/prodata/1/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq.*
- 7: /cgn2_6/prodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.6	3.3	1293	7	US-60-551-161-131
2	39.2	3.1	99971	6	US-10-767-471-10561
3	38.2	3.1	2232	6	US-10-808-727-45
4	37.4	3.0	110608	6	US-10-775-169-193
5	36.4	2.9	753	5	US-09-570-737A-647
6	35.8	2.9	765267	6	US-10-767-471-10616
7	35.4	2.8	2160	7	US-60-551-121-143
8	35.2	2.8	1184710	6	US-10-796-280-12394
9	35	2.8	1252	6	US-10-786-892-69
10	35	2.8	1790242	6	US-10-767-471-10805
11	34.8	2.8	1984	6	US-10-100-683-128
12	34	2.7	2427	5	US-09-537-710P-5
13	33.8	2.7	46143	6	US-10-796-280-12192
14	33.8	2.7	185248	6	US-10-796-280-12228
15	33.6	2.7	65385	6	US-10-796-307-8817
16	33.4	2.7	837	6	US-10-488-619-1789
17	33.4	2.7	158417	6	US-10-796-307-8778
18	33.4	2.7	158417	6	US-10-803-180-1675
19	33.2	2.7	65459	7	US-60-563-440-12039
20	33	2.7	276	6	US-10-724-972A-250
21	33	2.7	82553	6	US-10-100-683-12035
22	33	2.7	87394	1	PCT-US04-09319-6
23	32.8	2.6	201	6	US-10-796-280-21377
24	32.8	2.6	201	6	US-10-796-280-21644
25	32.8	2.6	142613	6	US-10-796-307-8749
26	32.8	2.6	154750	7	US-60-563-440-12165

27	32.8	2.6	663098	6	US-10-796-280-12289	Sequence 12289, A	
C	28	32.6	2.6	2610	6	US-10-784-480-243	Sequence 243, App
	29	32.6	2.6	3573	7	US-60-559-949-798	Sequence 798, App
C	30	32.6	2.6	31127	6	US-10-100-683-12644	Sequence 12644, A
	31	32.6	2.6	34655	6	US-10-796-280-12337	Sequence 12337, A
C	32	32.6	2.6	95990	7	US-60-563-440-11855	Sequence 11855, A
	33	32.6	2.6	163708	6	US-10-796-280-12588	Sequence 12588, A
C	34	32.6	2.6	546025	6	US-10-767-471-10717	Sequence 10717, A
	35	32.4	2.6	665	6	US-10-767-701-21191	Sequence 21191, A
C	36	32.4	2.6	643250	6	US-10-796-280-12230	Sequence 12230, A
	37	32.2	2.6	201	7	US-60-550-051-6793	Sequence 6793, Ap
C	38	32.2	2.6	359	5	US-09-796-692B-4345	Sequence 4345, Ap
	39	32.2	2.6	966	6	US-10-767-701-6462	Sequence 6462, Ap
C	40	32.2	2.6	61009	7	US-60-563-440-11881	Sequence 11881, A
	41	32.2	2.6	816609	7	US-60-550-051-2987	Sequence 2987, Ap
C	42	32	2.6	201	6	US-10-796-280-49577	Sequence 49577, A
	43	32	2.6	201	6	US-10-796-280-65927	Sequence 65927, A
C	44	32	2.6	477	6	US-10-767-701-28416	Sequence 28416, A
	45	32	2.6	531	6	US-10-767-701-29142	Sequence 29142, A

ALIGNMENTS

RESULT 1

US-60-551-161-131
; Sequence 131, Application US/60551161

; GENERAL INFORMATION:

; APPLICANT: Kluenhammer, Todd R.

; APPLICANT: Russell, William M.

; APPLICANT: Alterman, Eric

; APPLICANT: Cano, Raul J.

; APPLICANT: Hamrick, Alice

; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Stress-Related Proteins and Uses

; TITLE OF INVENTION: Therefore

; FILE REFERENCE: 043556/261831

; CURRENT APPLICATION NUMBER: US/60/551,161

; CURRENT FILING DATE: 2004-03-08

; NUMBER OF SEQ ID NOS: 282

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 131

; LENGTH: 1293

; TYPE: DNA

; ORGANISM: Lactobacillus acidophilus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(1293)

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (0)...(0)

; OTHER INFORMATION: ORF 1295; FtsY - Cell division protein

US-60-551-161-131

Query Match 3.3%; Score 41.6; DB 7; Length 1293;

Best Local Similarity 51.8%; Pred. No. 0.074; Mismatches 94; Indels 0; Gaps 0;

Matches 98; Conservative 0;

QY 437 ATAAAGAAAACCAATACAAAATCAAGCGAAATCTGTGTGATGTGGAATATCTGTACAGT 496

Db 59 AAAAGATGAGAAAACAAAGTCAAGACGAGAGATATAGAAAGAACCACTGATATGG 118

QY 497 TTTACTCCCTTAAACCTGTGATGAGATTCAGACAGTCTCAAGATCTAAAGTATTGA 556

Db 119 AAAATGACTCTTCTAATGTTGAGATTTCAGAAAAGTAGATATGATCTCAAGATTCGG 178

QY 557 AAACATGCTATCGGTGACACCATCTCAAGATTTACTAGCTCAAGCAACAAGCA 616

Db 179 AAAAATAGAGAAACTGAGAAAACACTTCTGATGATCGTTGAGAGAAATTA 238

QY 617 TTTTAAACAAA 628

Db 239 GTGAAATGATA 250

RESULT 2
 US-10-767-471-10561/c
 ; Sequence 10561, Application US/10767471
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001505
 ; CURRENT APPLICATION NUMBER: US/10/767,471
 ; PRIOR FILING DATE: 2004-01-30
 ; NUMBER OF SEQ ID NOS: 50231
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 10561
 ; LENGTH: 99971
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(99971)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
 ; US-10-767-471-10561
 Query Match 3.1%; Score 39.2; DB 6; Length 99971;
 Best Local Similarity 48.2%; Pred. No. 2.1;
 Matches 107; Conservative 1; Mismatches 114; Indels 0; Gaps 0;
 QY 674 CTCATGCAATGACATTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCATG 733
 DB 42609 CTTAAATCAGTGAACCTTACCACATTTTAGATTAAAGAAAAATATCATATGACCATTT 42550
 QY 734 TCAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGTCTGAATGAAGAAATAA 793
 DB 42549 AAAAGATTGAGGACAGATTTTATATAAATTCACATTTTATGTTGATTAAAGATAA 42490
 QY 794 ACAACACTGACCTGATCTCTGAGAAATATATAGCTCTTAAAAAGGGGAAAAAGCGGTATG 853
 DB 42489 TAACAAAACCTCTTAGCAAGCTTAGAGGATACCTCCATAATCTGGAAGGATACCTTCTA 42430
 QY 854 ATCCCTTTGATCGGAGTCACTTGAACCTGTTACCATCAAT 895
 DB 42429 AAACATCTATGCAATAAATCAATGAACCAAGTCATCTGAGAT 42389
 RESULT 3
 US-10-808-727-45
 ; Sequence 45, Application US/10808727
 ; GENERAL INFORMATION:
 ; APPLICANT: Oh, Steven
 ; APPLICANT: Chishtil, Athar
 ; APPLICANT: Liu, David
 ; APPLICANT: Li, Xuerong
 ; APPLICANT: Goel, Vikas
 ; TITLE OF INVENTION: Band 3-Binding Malaria Polypeptides and Uses Thereof
 ; FILE REFERENCE: S1237.7019US01
 ; CURRENT APPLICATION NUMBER: US/10/808,727
 ; PRIOR FILING DATE: 2004-03-25
 ; PRIOR APPLICATION NUMBER: US 10/087,464
 ; PRIOR FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 60/272,930
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 59
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 45
 ; LENGTH: 2232
 ; TYPE: DNA
 ; ORGANISM: Plasmodium falciparum
 ; US-10-808-727-45
 Query Match 3.1%; Score 38.2; DB 6; Length 2232;
 Best Local Similarity 51.1%; Pred. No. 0.86;
 Matches 116; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

QY 723 TACTTACCATTGTCAAAAAATCGGGAACAGCTTATGAGATCAATAAAAAATCTGGTCTGAA 782
 DB 1620 TAATCATCTCTGAAAAATTTATTTTAAAGAGAATATTTATGATGAAATGATGATATGGA 1679
 QY 783 TGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAGCTCTTAAAAAGGGGA 842
 DB 1680 AGTAAAAAGTTAAAAAATAGGTGTCACTTAAAAAATTTGAACCACTTAAAAATGG--- 1736
 QY 843 AAAGCCGTATGATCCCTTTTGATCGCAGTCACTTGAACCTGTTCCACCATCAAAATACGTTGA 902
 DB 1737 AAATGTTAGTGAACCACTTAAATTTGATTCTATTAGGAAATAAGATAAATAACACATTGA 1796
 QY 903 TGTCAACACCAACGAATTTGCTAAAAAGCGAGAGCTCTTTAACAGCTA 949
 DB 1797 AGCTATAAACCAAGCATATTCAAATTTATTAACCAAGAAATTACAAGCTA 1843
 RESULT 4
 US-10-775-169-193/c
 ; Sequence 193, Application US/10775169
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Burczynski, Michael
 ; APPLICANT: Twine, Natalie
 ; APPLICANT: Dörner, Andrew
 ; APPLICANT: Trepicchio, William
 ; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
 ; FILE REFERENCE: AM101080 (031896-013000)
 ; CURRENT APPLICATION NUMBER: US/10/775,169
 ; CURRENT FILING DATE: 2004-02-11
 ; NUMBER OF SEQ ID NOS: 5278
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 193
 ; LENGTH: 110608
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-775-169-193
 Query Match 3.0%; Score 37.4; DB 6; Length 110608;
 Best Local Similarity 52.2%; Pred. No. 7.3;
 Matches 83; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
 QY 415 GGACATGTGGCGTTAGACCATATAAGAAAAACCAATACAAATCAAGCGAATCTGTT 474
 DB 46041 GGGGATGGGAGCGAGAGTCTCTTCAAAAAATAAATAAACAACAAAAACAAATGAGAA 45982
 QY 475 GATGTGGAATATATCTGTACAGTCTTACTCCCTTAAACCTGTATGACGATTTTCAGACCCAGGT 534
 DB 45981 ACTGTAATAATAAACCGTAACCTGTGTGAATAGTATTAGGGAATCTCCACTAGAGT 45922
 QY 535 CTCAAAGATATCTAAGCTATTGAAACAACTAGCTATCGGT 573
 DB 45921 CTCTGAGTTTGAAGGTTTGTGAAATTTGAGTGTCTTCTGT 45883
 RESULT 5
 US-09-570-737A-647/c
 ; Sequence 647, Application US/09570737A
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqi
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedrick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Li, Samuel
 ; APPLICANT: Wang, Aijun

QY	746	ACAAGCTTATGAGATCAATAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACC	805
Db	440	AATACCCCTAATAACGCCAGATAAAACCCCAATTTTAAAGATTTTTTAAACTTTTACGCC	361
QY	806	TGATCTCTCAGAAATATTACGTCCTTTAAAAAAGGGGAAAGCCGATGATCCCTTTGATC	865
Db	380	TTGAATGCAGAAAAATCTATTTTCTATGAATAGGAAACCCGTAATAAACTTAAAGAAA	321
QY	866	GCAGTCACCTTGAAACTGTGTCACCATCAAAATACGTTGATGTCAACACCAACGAATTGCTA	924
Db	320	GGATAAAAAATAAAGAAGACAAACACGCTGAAGTCAAAACCAACAAATGGCTA	262

RESULT 10
 US-10-767-471-10805
 ; Sequence 10805, Application US/10767471
 ; GENERAL INFORMATION:
 ; APPLICANT: MICHÈLE et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001505
 ; CURRENT APPLICATION NUMBER: US/10/767,471
 ; CURRENT FILING DATE: 2004-01-30
 ; NUMBER OF SEQ ID NOS: 50231
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 10805
 ; LENGTH: 1790242
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)_(1790242)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
 US-10-767-471-10805

Query Match	2.8%	Score 35;	DB 6;	Length 1790242;
Best Local Similarity	48.6%	Pred. No. 99;		
Matches 122;	Conservative 1;	Mismatches 126;	Indels 2;	Gaps 1;

QY	249	TCAAGAACAAATTGATCGCTTAACTGTCACAGTAACGAGCACTACTTTGAGTCAATGATTT	308
Db	545236	TAAACARCAAGTTTAAAGAATAAATCTAAAGCAATTTGAAAATATTTTGAGACTAATGAAAT	545295
QY	309	TGCAACGCGATGCAACCACTTACTGATCGAAACGGCAAGTCTACTTTGCTGCACAAAGATGG	368
Db	545296	TGAAAACAAGCACACCAAGCTCATGGA--TGCAGCAAAACCACTACTAAGAGGGAAT	545353
QY	369	TTGCGTAACTTGGCCGACCCCACTGTCACAAAGATTTTGGCTAAGCGGACATGTGCGCGT	428
Db	545354	TTCCATGACTGTTTGACATTCAAAAGAAAGAAAGATTTTAAAGTAAATAATCTAACTTTGCA	545413
QY	429	TAGACCATATAAAGAAAACCAATACAAATCAAGCGAAATCTGTTGATCTGGAAATATAC	488
Db	545414	TCTCAATAAGTAGAAGAAAAGAGAACAAATTAAGCCCAAGATTTAGTAGAAGGAAAGAAAT	545473
QY	489	TGTACAGTTTA 499	
Db	545474	GGTAAGATT 545484	

RESULT 11
 US-10-100-683-128/c
 ; Sequence 128, Application US/10100683
 ; GENERAL INFORMATION:
 ; APPLICANT: ROSEN, et al.
 ; TITLE OF INVENTION: Human Secreted Proteins
 ; FILE REFERENCE: PS900
 ; CURRENT APPLICATION NUMBER: US/10/100,683
 ; CURRENT FILING DATE: 2002-03-19
 ; PRIOR APPLICATION NUMBER: US 60/040,162
 ; PRIOR FILING DATE: 1997-03-07
 ; PRIOR APPLICATION NUMBER: US 60/043,576

```

; OTHER INFORMATION: n= a or g or c or t/u
US-09-537-710D-5

Query Match      2.7%; Score 34; DB 5; Length 2427;
Best Local Similarity 50.6%; Pred. No. 14;
Matches 82; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 205 GCGATGCCACATAAACTTCGAAAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTGATC 264
DB 1357 GCTGCGCCCATCTGAGATTAAATAATTGATTTTCGAGGTGCTGTCAAAGGTCAAGATATC 1416
QY 265 GCTAACGTCACAGTAACGACGACTACTTTTGAGGTCAATTGATTTTGCAGCGATGCAACC 324
DB 1417 CCAATCACACCTGTGCTGACGTGTGGACAGAGTACCATGACATGGAAATTGCTGGGATC 1476
QY 325 ATTACTGATCGAAAACGGCAAGCTCTACTTTTGTCTGCAAAAGAT 366
DB 1477 AAAGCTATCGCTGAGTATAAGGCTACACTGCTGCTGGAAGCT 1518

RESULT 13
US-10-796-280-12192/c
; Sequence 12192, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12192
; LENGTH: 46143
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-796-280-12192

Query Match      2.7%; Score 33.8; DB 6; Length 46143;
Best Local Similarity 48.2%; Pred. No. 54;
Matches 95; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 453 ACAAATCAAGCGAATCTGTTGATGTGGAATATATCTGACAGTTTACTCCCTTAAACCC 512
DB 19113 AGAAGATTTTGAGAGTATTCTGACATTCATTTGTAAACAAGGCAATTTGTAAACAAT 19054
QY 513 TGATGACGATTTTCAGACCAAGGCTCTCAAGATACTAAGCTATTGAAAACACTAGCTATCGG 572
DB 19053 TGCAAAGATTTGGTAAACAATTCGAAACGATTTGTAACAATTTGCAACAACTGTGTAACA 18994
QY 573 TGACACATTCACATCTCAAGAAATTACTAGCTCAAGCACAAAGCATTTTAAACAAACCCA 632
DB 18993 ACTTGTAAACAATTTGCAACAATTTGTTAAATAAGTTGTTAAACAATTCACAAACAAAGCCCA 18934
QY 633 CCCAGGCTATACGATTT 649
DB 18933 GAGAGGGGAAGTAATT 18917

RESULT 14
US-10-796-280-12228
; Sequence 12228, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12228
; LENGTH: 185248

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 6, 2004, 01:11:57 ; Search time 3628 Seconds
(without alignments)
10247.634 Million cell updates/sec

Title: US-09-882-509-3

Perfect score: 1245

Sequence: 1 attgttgacctgagtgct.....ataaccctaagacaaataa 1245

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estro:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rtd:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	51.6	4.1	712	13 BX416727	BX416727 BX416727
c 2	49.4	4.0	1124	13 BX436282	BX436282 BX436282
c 3	49.4	4.0	1217	13 BQ678748	BQ678748 AGENCOURT
c 4	48	3.9	1175	13 BX423299	BX423299 BX423299

c 5	47.8	3.8	1101	29 CNS0100X	AL098379 Drosophil
c 6	47	3.8	994	13 BX414650	BX414650 BX414650
c 7	46.4	3.7	1101	29 CNS0039G	AL063921 Drosophil
c 8	46.2	3.7	885	13 BX425603	BX425603 BX425603
c 9	45.4	3.6	1085	13 BX414831	BX414831 BX414831
c 10	45.4	3.6	1201	13 BX361080	BX361080 BX361080
c 11	44.6	3.6	1007	29 CNS06X9S	AL078462 T3 end of
c 12	44.4	3.6	1101	29 CNS00LT2	AL078714 Drosophil
c 13	44.2	3.6	878	29 CNS0187R	AL108993 Drosophil
c 14	44.2	3.6	1101	29 CNS00HE9	AL073374 Drosophil
c 15	44.2	3.6	1109	13 BX375784	BX375784 BX375784
c 16	42.8	3.4	1101	29 CNS00LOO	AL068607 Drosophil
c 17	42.4	3.4	1180	13 BX436369	BX436369 BX436369
c 18	42	3.4	1061	13 BX414576	BX414576 BX414576
c 19	41.8	3.4	500	12 BP188195	BP188195 BP188195
c 20	41.8	3.4	1201	13 BX419736	BX419736 BX419736
c 21	41.6	3.3	990	13 BX382345	BX382345 BX382345
c 22	41.6	3.3	1081	28 CC238890	CC238890 CH261-80N
c 23	41	3.3	796	12 BM398641	BM398641 5009-0-48
c 24	40.8	3.3	508	28 AQ591034	AQ591034 HS 5391 B
c 25	40.8	3.3	678	28 BH473995	BH473995 BQNW637R
c 26	40.8	3.3	741	14 CF134481	CF134481 CFG12-85
c 27	40.8	3.3	976	29 AG136173	AG136173 Pan trogl
c 28	40.8	3.3	1101	29 CNS000D1	AL065414 Drosophil
c 29	40.8	3.3	1103	13 BX387723	BX387723 BX387723
c 30	40.8	3.3	1201	13 BX394291	BX394291 BX394291
c 31	40.6	3.3	507	12 BM679961	BM679961 ku03c03 Y
c 32	40.6	3.3	996	29 CNS00FUH	AL071063 Drosophil
c 33	40.6	3.3	1091	13 BX424950	BX424950 BX424950
c 34	40.2	3.2	337	29 AL765028	AL765028 Arabidops
c 35	40.2	3.2	500	9 AU088371	AU088371 AU088371
c 36	40	3.2	649	9 AV733574	AV733574 AV733574
c 37	40	3.2	895	13 BQ729093	BQ729093 AGENCOURT
c 38	40	3.2	1101	29 CNS00ESI	AL069797 Drosophil
c 39	40	3.2	1412	28 BZ569852	BZ569852 msh2_1049
c 40	39.8	3.2	828	28 ABZ686097	ABZ686097 ENTLI59TF
c 41	39.6	3.2	780	10 BE642964	BE642964 Cri12 7 IO
c 42	39.6	3.2	1024	28 CC437374	CC437374 PUNW25TD
c 43	39.4	3.2	387	14 CA859923	CA859923 ESTG3417
c 44	39.4	3.2	646	9 AL045329	AL045329 DXFZp434A
c 45	39.4	3.2	715	28 BH924586	BH924586 odi55e06

ALIGNMENTS

RESULT 1	BX416727	712 bp	mRNA	linear	EST 15-MAY-2003
BX416727/c	BX416727	Homo sapiens	NEUROBLASTOMA	Homo sapiens	cdna clone
LOCUS	CS0DA011Y114	5-PRIME	mRNA sequence.		
DEFINITION	BX416727	EST.			
ACCESSION	BX416727.1	GI:30765629			
VERSION	BX416727.1	GI:30765629			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 712)				
AUTHORS	Li.W.B., Gruber,C., Jessee,J. and Polayes,D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Genoscope				
	Genoscope - Centre National de Sequencage				
	BP 191 91006 EVRY cedex - France				
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr				
	Library was constructed by Life Technologies, a division of				
	Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :				
	http://fulllength.invitrogen.com/ InvitroGen Corporation 1600				
	Faraday Avenue Genoscope sequence ID : CS0DA011BE07Q1.				
	Location/Qualifiers				
	1. .712				
	/organism="Homo sapiens"				

```
/mol_type="mrna"  
/db_xref="taxon:9606"  
/clone="CS0DA011Y14"  
/tissue_type="NEUROBLASTOMA"  
/clone_lib="Homo sapiens NEUROBLASTOMA"  
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."
```

ORIGIN

```
Query Match 4.1%; Score 51.6; DB 13; Length 712;  
Best Local Similarity 12.1%; Pred. No. 0.06;  
Matches 49; Conservative 180; Mismatches 175; Indels 0; Gaps 0;  
  
QY 432 ACCATATAGAGAAAACCAATACAAATCAAGCAATCTGTTGATGTGGAATATCTGT 491  
Db MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  
650  
  
QY 492 ACAGTTTACTCCCTTAAACCTGATGAGATTTCAGACCGGTCTCAAGAGATCTAAGCT 551  
Db MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  
590  
  
QY 552 ATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTTCTAGCTCAAGCACA 611  
Db MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  
530  
  
QY 612 AAGCATTTTAAACAAACCCAGGCTATACATTTATGAACGCTGACTCTCAATCGT 671  
Db MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  
470  
  
QY 672 CACTCATGCAATGACATTTTCCGTAGCATTTTACCAATGATGATCAAGATTCTTACCA 731  
Db MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  
410  
  
QY 732 TGTCAAAATCGGGAACAGCTTATGAGATCAATAAATAATCTGGTCTGAATGAAGAAAT 791  
Db MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN  
350  
  
QY 792 AAACAACACTGACCTGCTCTGAGAAATATTACGCTCTTAA 835  
Db AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 306
```

```
RESULT 2  
BX436282/c 1124 bp mRNA linear EST 15-MAY-2003  
LOCUS BX436282 Homo sapiens THYMUS cDNA clone CS0CAP001YC01  
DEFINITION S-PRIME, mRNA sequence.  
ACCESSION BX436282  
VERSION BX436282.1 GI:30787521  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1124)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0CAP001AB01Q1.  
Location/Qualifiers  
1. 1124  
/organism="Homo sapiens"  
/mol_type="mrna"
```

```
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT
```

FEATURES

```
source  
1. 1124  
/organism="Homo sapiens"  
/mol_type="mrna"
```

```
/db_xref="taxon:9606"  
/clone="CS0CAP001YC01"  
/tissue_type="THYMUS"  
/clone_lib="Homo sapiens THYMUS"  
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."
```

ORIGIN

```
Query Match 4.0%; Score 49.4; DB 13; Length 1124;  
Best Local Similarity 22.2%; Pred. No. 0.21;  
Matches 129; Conservative 174; Mismatches 278; Indels 0; Gaps 0;  
  
QY 492 ACAGTTTACTCCCTTAAACCTGATGAGATTTCAGACCGGTCTCAAGATCTAAGCT 551  
Db AYYYYYYYYYYYYYAAAAAAAAAAAAAAAAHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
1008 AYYYYYYYYYYYYYAAAAAAAAAAAAAAAAHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
949  
  
QY 552 ATTGAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTTCTAGCTCAAGCACA 611  
Db YYYHHAAAYYYYYYYYYYYYYYVAYYHHAAHHHHHHHHHHHHHHHHHHHHHHHHHH  
948 YYYHHAAAYYYYYYYYYYYYYYVAYYHHAAHHHHHHHHHHHHHHHHHHHHHHHHHH  
889  
  
QY 612 AAGCATTTTAAACAAACCCAGGCTATGATTTATGAACGCTGACTCTCAATCGT 671  
Db AYYYYYYYYYYYYYHHAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
888 AYYYYYYYYYYYYYHHAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
829  
  
QY 672 CACTCATGCAATGACATTTTCCGTAGCATTTTACCAATGATGATCAAGATTCTTACCA 731  
Db YAAAAAAAAAAAAAAAAAYYYYYYYYYYAAAAAAAAAAAAAAAAHHHHHHHHHHHHHHHH  
828 YAAAAAAAAAAAAAAAAAYYYYYYYYYYAAAAAAAAAAAAAAAAHHHHHHHHHHHHHHHH  
769  
  
QY 732 TGTCAAAATCGGGAACAGCTTATGAGATCAATAAATAATCTGGTCTGAATGAAGAAAT 791  
Db YYYYYYAAHHVHAHHAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
768 YYYYYYAAHHVHAHHAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
709  
  
QY 792 AAACAACACTGACCTGCTCTGAGAAATATTAGCTCTTAAATAAGGGGAAAGCCGTA 851  
Db AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
708 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
649  
  
QY 852 TGATCCCTTGTGCGAGTCATCTGAACTGTTCCATCAATACGATGATGTCACAC 911  
Db AHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
648 AHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
589  
  
QY 912 CAACGAATTCATAAAGCGAGCAGCTCTTAAACAGCTAGCGAACGTAACCTAGCTTCAG 971  
Db YHAMYYYYYYYYYHHAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
588 YHAMYYYYYYYYYHHAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
529  
  
QY 972 AGATTTATAGCTCCTCGTATGAGGTAACCTACTCTACAACAATCTCGATGCTTTGG 1031  
Db AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
469  
  
QY 1032 TATTATGAGCTATACCTTAACTGGAAGAGTAGAGGATAATC 1072  
Db YAAAAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
468 YAAAAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH  
428
```

```
RESULT 3  
BQ678748  
LOCUS BQ678748 1217 bp mRNA linear EST 15-JUL-2002  
DEFINITION AGENCOURT_8241036 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6265192  
S', mRNA sequence.  
ACCESSION BQ678748  
VERSION BQ678748.1 GI:21791427  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1217)  
NIH-MGC http://mgs.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.
```

Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2431 row: o column: 17
High quality sequence step: 143.
Location/Qualifiers
1. 1217
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6265192"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 112"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GSCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."

BX423299 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
 CS0DF08YR06 3-PRIME, mRNA sequence.
 BX423299
 BX423299.1 GI:30655486
 EST.
 Homo sapiens
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 1175)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8830.r For
more information about this cluster, see
<http://www.genoscope.cns.fr/>
cgi-bin/cluster.cgi?seq=CS1AF002ZH04NP1&cluster=8830.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1AF002ZH04NP1.

FEATURES

ORIGIN

Query Match	4.0%	Score	49.4	DB	13	Length	1217
Best Local Similarity	43.7%	Pred.	No.	0.22			
Matches	218	Conservative	0	Mismatches	281	Indels	0
QY	447	ACCAATACAAAAATCAAGCGAAATCTGTTGATGTCGATATACTGTACAGTTTACTCCCTT	506				
Db	282	AACACGACAAACCAAGGCAAAACAAAGAAAAACCAACAAAAAACCACCCACCGA	341				
QY	507	AAACCCCTGATGACGATTTTCAGACGAGTCTCAAGATACTTAAGCTATTGAAAAACACTAGC	566				
Db	342	AAAGGAAAGGAAAAATAACTAACCCGACGAAGAAGCAAAAATAACAAAAAANAANAANA	401				
QY	567	TATCGGTGACACATCATCTCAAGATTTATAGTCTCAAGCAAAAGCAATTTTAAACAA	626				
Db	402	CAAAGGAAACACCGCCACCAAGAAAAACCAACCAACCAAGAAACAAACCAATAATAACAC	461				
QY	627	AACCCACCCAGGCTATACGATTTATGAACGTGACTCTCAATCGTCACTCATGACATGA	686				
Db	462	AACAAACGAACAGAAACAGCAGAGAGACAACCGACAATAACCAATATACAAACAAAGA	521				
QY	687	CATTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGA	746				
Db	522	AAAAAANAANAATACGAAANAATAGAGAAAAAAGAAAAAGAAATAAANAANAATAAGGA	581				
QY	747	ACNAGCTTATGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACCAACTGACCT	806				
Db	582	AAAAAGAACAAAGAACCAACGAGAAAAATACAAAAAAGAAATAGAAAAAACAANAACACA	641				
QY	807	GATCTCTGAGAAATATTACGTCTCTTAAAAAGGGGAAAGCCGTATGATCCCTTTGATCG	866				
Db	642	ACNCCCAACAAACCGACATGATGAATTAACCGGAATCGAAACCAATCAATATAACAC	701				
QY	867	CAGTCACTTTGAAACTGTTTCAACCAATCAAAATACGTTTGATGTCAACCAACGAATTGCTAAA	926				
Db	702	AAAAACAACAAAAACAACAACAAACAGAAACGAGAGATATAACATCAACAAATACAA	761				
QY	927	AAGCGAGCGCTCTTAAACA	945				
Db	762	TCAACCAACGACAAAAATA	780				

RESULT 4
BX423299/c
LOCUS

Query Match	3.9%;	Score 48;	DB 13;	Length 117/5;
Best Local Similarity	36.2%;	Pred. No. 0.48;		
Matches 117;	Conservative 58;	Mismatches 145;	Indels 3;	Gaps 1;
581	TCACATCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCACCCAGGCT	640		
1098	TANYTTTTTCMAAAATWAAAATVTTAWAWATAAARCAAAATAAAAAARCCCVYAWWC	1039		
641	ATACGATTATGAAGCGTGACTCCTCAATCGTCACTCATGACAAATGACATCTTTCCGTAC	700		
1038	YATTATTATTAWTBMAHAWATATTATTAWMAAAWATATAAAAAAATAAATAATTAHT	979		
701	TTTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAATCGGGAAACAAGCTTTATG	760		
978	WTTTWAAMAAAAATTAAATTTWTTTAAATVTTVYAAWAAAAAAMWAAAAAAWTTAA	919		
761	TCAAATAAAAAAATCTGGTCTCGATGAAGAAATAACCAACACTGACCTGATCTCTGAG	820		
918	TTAAWAAAHWTAAAA--TAAATTTWAAAATAATAATVYAAAYAAAAAMWYYTCAAAAT	862		
821	ATTAGTCCTCTAAAAAAGGGAAAAAGCGGTATGATCCCTTTTGATCGCAGTCACTTTGA	880		
861	AAHWATCCATATHCAATCAATVYTTVATVYCCATCTCTVYAKCBMBTCTTTTMTWA	802		
881	TGTTCCACATCAAAATACGTTGAT	903		
801	TTAAWAAAAATTTWTAAKYTTTCT	779		

RESULT 5
CNS0100X/c
LOCUS

[illegible]

```

RESULT 9
BX414831/c
LOCUS
DEFINITION
1085 bp mRNA linear EST 15-MAY-2003
BX414831 Homo sapiens THYMUS Homo sapiens cdna CSOCAP002YL02
5'-PRIME, mRNA sequence..
ACCESSION
BX414831
VERSION
BX414831.1 GI:30767501
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
1 (bases 1 to 1085)
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Haraday Avenue Genoscope sequence ID : CSOCAP002F01QP1.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

FEATURES
source
1..1085
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOCAP002YL02"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
rayway Avenue geloscope sequence ID : CSOCAP002F01QPL1.
Location/Qualifiers

```

with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

```

ORIGIN
Query Match 3.6%; Score 45.4; DB 13; Length 1085;
Best Local Similarity 31.2%; Pred. No. 2.1;
Matches 104; Conservative 69; Mismatches 160; Indels 0; Gaps 0;

Qy 439 AAAGAAAACCAATACAAAATCAAGCGAAATCTGTTGATGTGGAATATATCTGTACAGTTT 498
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1066 AAAYAMAAACWAAAMWAAAYCCCHWCMMWAAATHYTAWAMYCWAMWTHATCTTCTWATMTY 1007
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 499 ACTCCCTTAAACCCCTGATGACGATTTTCAGACCAAGGTCCTCAAAGATACCTAAGCTATTGAAA 558
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1006 MYACYAAAMCMYCHNAMCCYMYCYHNAHCYCYTYAAAMAMCAAAAATATAMWAAAMH 947
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 559 ACCTAGCTATCGGTGACACCACTCACATCTCAAGAATTTACTAGCTCAAGCACAAAGCATT 618
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 946 MCAWAATACCCCMCTCCTAHAAATAWAAAAATAAMTTATYMMYTCYMYCYCWCTCTAACYC 887
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 619 TTAAAAAAAACCCACCCAGCGCTATACGATTTATGAAGCTGACTCTCTCAATCGTCACTCAT 678
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 886 ATAATTAAAAAHHAAAYVHHHTAMAHATHTCTMTWAAAAAMWYAYATAACCCCTTHTTTAAT 827
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 679 GACAATTGACATTTTCCGTACGATTTTACCATCGATCAAGAGTTTACTTACCACTGTCAAA 738
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 826 YTWAAAYAMCTTAAATTTTAAAYCTTYAAMTCTCMAAAYATAAAMCCACHACAMY 767
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 739 AATCGGGAACAAGCTTTATGAGATCAATAAAAAA 771
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 766 ATAYYAMAAAYAYCYXYAMTCCMWAACAAACMWA 734
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 10	
EX361080/c	
LOCUS	1201 bp mRNA linear EST 05-MAY-2003
DEFINITION	EX361080 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0D1079YN15 3-PRIME, mRNA sequence.
ACCESSION	EX361080
VERSION	EX361080.1 GI:30374504
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Li, W.-B., Gruber, C., Jesse, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen/Corporation 1600 Paradise Avenue Genoscope sequence ID : CS0D1079PG08NP1.

```

FEATURES
source
    1..1201
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CSOD1079YN16"
        /tissue_type="PLACENTA COT 25-NORMALIZED"
        /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
        /note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and Ecor V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

```

Query Match      3.6%; Score 45.4; DB 13; Length 1201;
Best Local Similarity 12.4%; Pred. No. 2.1;
Matches 27; Conservative 111; Mismatches 80; Indels 0; Gaps 0;

QY 427 GTTAGACCATATTAAGAAACCAACCAATACAAATCAAGCGAAATCTGTGTGATGCGAATAT 486
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 976 KKAAGAADAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 917
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 487 ACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTCAGACAGGTCTCAAGATACT 546
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 916 MVMVMVMVTTTTMMVMVMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMM 857
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 547 AAGCTATTGAAACACTAGTATCGGTGACACCATCACATCTCAAGAAATTAATAGTCTCA 606
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 856 MVMVMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMM 757
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 607 GCACAAAGCATTTTAACAAACCCACCCAGGCTATAC 644
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 796 KXETRYKKKKTTCMMVMVMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMM 759
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 11
CNS06X9S
LOCUS
DEFINITION
T3 end of clone AX00A039F08 of library AX00A from strain CBS 7064
AL419462
VERSION
AL419462.1 GI:12202640
KEYWORDS
GSS.
SOURCE
Pichia farinosa
ORGANISM
Pichia farinosa
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE
1 (bases 1 to 1007)
Souciet J.L., Aigle M., Artiguenave F., Blandin G.,
Bolotin-Fukuhara M., Bon E., Brottier P., Casaregola S.,
de-Montigny J., Dujon B., Durand P., Lepingle A., Llorente B.,
Malpertuy A., Neuveglise C., Ozier-Kalogeropoulos O., Potier S.,
Saurin W., Tekaia F., Toffano-Nioche C., Wesolowski-Louvel M.,
Wincker P. and Weissenbach J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)
20584711
11152876
2 (bases 1 to 1007)
de Montigny J., Spehner C., Souciet J., Tekaia F., Dujon B.,
Wincker P., Artiguenave F. and Potier S.
Genomic exploration of the hemiascomycetous yeasts: 15. Pichia
sorbitophila
FEBS Lett. 487 (1), 87-90 (2000)
20584725
11152890
3 (bases 1 to 1007)
Genoscope.
Direct Submission
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
Location/Qualifiers
1..1007
/organism="Pichia farinosa"
/mol_type="genomic DNA"

```

```

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
SOURCE
ORGANISM
ACCESSION
VERSION
AL078714.1 GI:5102004
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aron Mammos in Pictet de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA and was provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or

```

```

FEATURES
source

```

```

/strain="CBS 7064"
/db_xref="taxon:4920"
/clone="AX00A039F08"
/clone_lib="AX00A"
/note="end : T3"

```

```

Query Match      3.6%; Score 44.6; DB 29; Length 1007;
Best Local Similarity 35.0%; Pred. No. 3.2;
Matches 143; Conservative 51; Mismatches 215; Indels 0; Gaps 0;

QY 437 ATAAAGAAAACCAATACAAATCAAGCGAAATCTGTGTGATGCGAATATACATGTCAGT 496
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 AAWTAAAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 264
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 497 TTACTCCCTTAAACCCCTGATGACGATTCAGACAGGTCTCAAGATTAATTAAGCTATTGA 556
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 265 AAWTAAAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 324
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 557 AAACCTAGCTATCGGTGACACCATCACATCTCAAGAAATTAATTAAGCTCAAGCAACAAGCA 616
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 325 AAWTAAAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 384
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 617 TTTTAAACAAACCCACCCAGGCTATACGATTTATGACAGTCTCTCAATCGTCACTC 676
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 385 TGTAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 444
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 677 ATGCAATGACATTTTCCGTGATGATTTTACCAATGATCAAGAGTTTACTTACCATGTCA 736
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 445 GAAATAAAAGATTTAAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 504
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 737 AAAATCGGGAACAAGCTTATGATCAATAAAATCTGCTGATGAATGAATATAACA 796
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 505 AAAATAAAWATWTAACWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 564
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 797 ACACGTACCTGATCTCTGAGAATATTACGTCCTTAAAAAAGGGGAAAA 845
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 565 AAAATAAAWATAATACGATGATTATACACGTTTAGTGGGGGAATAA 613
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 12
CNS00LT2
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC:
BAC48P19 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL078714.1 GI:5102004
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aron Mammos in Pictet de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA and was provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or

```

```

COMMENT

```


/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BAC34A13"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN

```

Query Match      3.6%; Score 44.2; DB 29; Length 1101;
Best Local Similarity 32.7%; Pred. No. 4.1;
Matches 109; Conservative 7; Mismatches 217; Indels 0; Gaps 0;

QY 513 TGATCAGCATTCAGACAGGCTCTCAAGATCTACTAGCTATTGAAACACACTAGCTATCGG 572
Db 428 TTAKGWTATGGAATATGGTGGKAAATMTAAANAAATAAATAAAMNAAAAAACAAC 369
QY 573 TGACACCATCATCTCAAGATTAATCTAGCTCAAGCACAAGCAATTTTAAACAAACCCA 632
Db 368 ACAANNANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 309
QY 633 CCCAGGCTATAGATTATGACGCTGCTCTCAATCTGCTACTCATGACATGCAATTTT 692
Db 308 CGCANNANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 249
QY 693 CCGTACGATTTTACCAATGGATCAAGCTTTTACTTTACCATGTCRAAAATCGGACACAGC 752
Db 248 NNNAANNNNGGAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 189
QY 753 TTATGAGATCAATAAAAAATCTGGTCTGAATCAAGAAATAAACACACTGACCTGATCTC 812
Db 188 NAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 129
QY 813 TGAGAAATATTACGCTCTTAAAAAGGGGAAAA 845
Db 128 AGNNGAANNNNNANGNNAANNAANNAAGGAGGNANA 96

```

RESULT 15

```

LOCUS      BX375784
DEFINITION BX375784 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC017YF05 5-PRIME, mRNA sequence.
ACCESSION  BX375784
VERSION    BX375784.1 GI:30448443
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1109)
            Li, W. B., Gruber, C., Jessee, J., and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
COMMENT    Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 8559.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DC017CC03QP1&cluster=8559.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Paraday Avenue Genoscope sequence ID : CS0DC017CC03QP1.
            Location/Qualifiers
              1. .1109

```

FEATURES

```

source     /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DC017YF05"
            /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
            /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

```

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

```

Query Match      3.6%; Score 44.2; DB 13; Length 1109;
Best Local Similarity 34.8%; Pred. No. 4.1;
Matches 103; Conservative 31; Mismatches 162; Indels 0; Gaps 0;

QY 541 GATACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCACATCTCAAGAAATCTA 600
Db 625 GNTAAAAAANTGTTNNAANANNCNAAATANANNNAANNNNNAATTTNAAAAAANA 684
QY 601 GCTCAAGCACAAGCAATTTTAAACAAACCCACCCAGGCTATACGATTTATGAACGTGAC 660
Db 685 GGTNAAAGGGTNAAAAAATGNGNNAAAAAAANTKNNANTCNNTYNTTTTATNNNTCHAAT 744
QY 661 TCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATCGATCAAGAG 720
Db 745 TMYATTATYATRAAATTAWAATYCTCAYTMTYTAATYTTTMTWAAACAAATC 804
QY 721 TTTACTTACCATGTCAAAAAATCGGGAACAAGCTTTATGAGATCAATAAAAAATCTGGTCTG 780
Db 805 WTAATYCWAATYATYCYAAAYCAAAAAATCTCAAWAAAAAATCMAWATATTAATCATCTC 864
QY 781 AATCAAGAAATAACAACACACTGACCTGATCTCTGAGAAATATATTACGTCCTTTAAAAA 836
Db 865 CYATMTCAAAAAHHAAAAAAYCCCTTACTMAAAAAAATAAYAAATCTCTCAAM 920

```

Search completed: May 6, 2004, 04:55:21
Job time : 3634 secs